Chapter 4

The Analytic Hierarchy Process and Expert Choice

The Analytic Hierarchy Process

The analytic hierarchy process (AHP), developed at the Wharton School of Business by Thomas Saaty[1], allows decision makers to model a complex problem in a hierarchical structure showing the relationships of the goal, objectives (criteria), sub-objectives, and alternatives (See Figure 1). Uncertainties and other influencing factors can also be included.

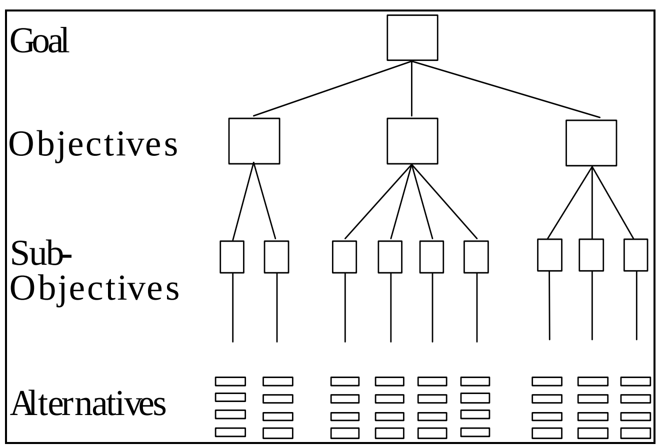


Figure 1 – Decision Hierarchy

AHP allows for the application of data, experience, insight, and intuition in a logical and thorough way. AHP enables decision-makers to *derive* ratio scale priorities or weights as opposed to arbitrarily *assigning* them. In so doing, AHP not only supports decision-makers by enabling them to structure complexity and exercise judgment, but allows them to incorporate both

---

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

objective and subjective considerations in the decision process[2]. AHP is a compensatory decision methodology because alternatives that are deficient with respect to one or more objectives can compensate by their performance with respect to other objectives. AHP is composed of several previously existing but unassociated concepts and techniques such as hierarchical structuring of complexity, pairwise comparisons, redundant judgments, an eigenvector[3] method for deriving weights, and consistency considerations. Although each of these concepts and techniques were useful in and of themselves, Saaty's synergistic combination of the concepts and techniques (along with some new developments) produced a process whose power is indeed far more than the sum of its parts.

With the introduction of its PC implementation, Expert Choice[4], the number and diversity of AHP applications has grown rapidly[5]. As of 1995, Expert Choice was being used in 57 countries throughout the world and there were over 1000 journal and other citations about AHP. The International Society of the Analytic Hierarchy Process[6] conducts meetings every two or three years throughout the world (the first meeting was held in Tianjin, China).

The ability for AHP to enhance the (evaluation and) choice phase of decision-making is well known. What is not as well known, however, is AHP's utility in *any* facet of problem solving and decision-making that involves evaluation and measurement. Forecasting is one such area. In the process of evaluating the alternatives to a decision, it often becomes apparent that the outcomes of one or more of alternative courses of action are uncertain. AHP can be used to measure the relative impact of numerous influencing factors on the possible outcomes and, in so doing, forecast (derive the distribution of relative likelihoods of) outcomes. These forecasts are then used when evaluating the alternative courses of action. Another

---

[2] Forman, E.H., "The Analytic Hierarchy Process as a Decision Support System," *Proceedings of the IEEE Computer Society* (Fall, 1983).
[3] Eigenvectors will be discussed later.
[4] Ernest H. Forman, Thomas L, Saaty, Mary Ann Selly, Rozann Waldron, *Expert Choice*, Decision Support Software, McLean, VA, 1983.
[5] F. Zahedi, The Analytic Hierarchy Process- A Survey of the Method and its Applications. *Interfaces*, (Vol. 16, 1986), pp. 96-108.
[6] Located on the Internet at http://ahp.net/www/ahp/support/ahp_disc.html

area of application of AHP beyond the choice phase of decision-making is in resource allocation. These, as well as other applications, will be addressed in this book.

Beyond Weights and Scores

The Analytic Hierarchy Process overcomes the problems with weights and scores approaches discussed above. This is done by structuring complexity as a hierarchy and by deriving ratio scale measures through pairwise *relative* comparisons. The pairwise comparison process can be performed using words, numbers, or graphical bars, and typically incorporates redundancy, which results in a reduction of measurement error as well as producing a measure of consistency of the comparison judgments.

Humans are much more capable of making relative rather than absolute judgments. The use of redundancy permits accurate priorities to be derived from verbal judgments even though the words themselves are not very accurate[7]. This opens up a new world of possibilities—we can use words to compare qualitative factors and derive ratio scale priorities that can be combined with quantitative factors!

Weights or priorities are not arbitrarily "assigned".

By using the AHP pairwise comparison process, weights or priorities are derived from a set of judgments[8]. While it is difficult to justify weights that are arbitrarily assigned, it is relatively easy to justify judgments and the basis (hard data, knowledge, experience) for the judgments. These weights or priorities are ratio level measures[9], not counts. In a Wall Street Journal article, "We Need to Measure, Not Count", Peter Drucker emphasized the need for measuring as opposed to counting[10]:

---

[7] Expert Choice also has a numerical mode, which, for numerical aspects of a problem would be even more "accurate". But it is not always appropriate to use numbers in such a direct fashion because priorities derived directly from accurately measured factors do not take into account the decision makers utility!
[8] Expressed either verbally, numerically, or graphically.
[9] See pages 34.
[10] Peter F. Drucker, "We Need to Measure, Not Count", *The Wall Street Journal Tuesday*, April 13, 1993.

"Quantification has been the rage in business and economics these past 50 years. Accountants have proliferated as fast as lawyers. Yet we do not have the measurements we need.

Neither our concepts nor our tools are adequate for the control of operations, or for managerial control. And, so far, there are neither the concepts nor the tools for business control - i.e., for economic decision-making. In the past few years, however, we have become increasingly aware of the need for such measurements…".

Furthermore, the priorities that are derived from judgments automatically incorporate the necessary non-linearities in measuring utility. For example, when considering a vehicle for city driving, the preference for a vehicle with a top speed of 40 miles per hour is probably more than twice that of a vehicle with a top speed of 20 miles per hour. But the preference for a vehicle with a top speed of 100 miles per hour would be much less than twice as preferable than a vehicle with a top speed of 50 miles per hour. When organizations lack the ability to measure, including non-linearities and utilities, they sometimes resort to counting as a poor substitute. "How many articles do I need for tenure?", a junior faculty member often asks. Depending on the contribution, perhaps just one, a wise Full Professor answers! The Analytic Hierarchy Process provides the ability to measure.

Inconsistency

The theory of AHP does <u>not</u> demand perfect consistency.

AHP allows inconsistency, but provides a measure of the inconsistency in each set of judgments. This measure is an important by-product of the process of deriving priorities based on pairwise comparisons. It is natural for people to want to be consistent. Being consistent is often thought of as a prerequisite to clear thinking. However, the real world is hardly ever perfectly consistent and we can learn new things only by allowing for some inconsistency with what we already know.

If we are perfectly consistent (as measured with an AHP inconsistency ratio of zero), we can <u>not</u> say that our judgments are good, just as we can not say that there is nothing wrong with us physically if our body temperature is 98.6 degrees. On the other hand, if our inconsistency is say 40 or 50% (an inconsistency ratio of 100% is equivalent to random judgments), we can say there is something wrong, just as we can say that there is something wrong if our body temperature is 104 degrees.

An inconsistency ratio of about 10% or less is usually considered "acceptable", but the particular circumstance may warrant the acceptance of a higher value[11]. Let us look at some of the reasons why inconsistency occurs as well as the useful information that the inconsistency ratio conveys, and ways to reduce it.

Causes of Inconsistency

Clerical Error

The most common cause of inconsistency is a clerical error. When entering one or more judgments into a computer, the wrong value, or perhaps the inverse of what was intended is entered. Clerical errors can be very detrimental and often go undetected in many computer analyses[12]. When using Expert Choice, one can easily find and correct such errors.

Lack of Information

A second cause of inconsistency is lack of information. .If one has little or no information about the factors being compared, then judgments will appear to be random and a high inconsistency ratio will result[13]. Sometimes we fool ourselves into thinking that we know more than we really do. It is useful to find out that a lack of information exists, although sometimes we might be willing to proceed without immediately spending time and money gathering additional information in order to ascertain if the additional information is likely to have a significant impact on the decision.

---

[11] For example, a body temperature of 100 degrees may be taken as normal if we know that the person has just completed a 26 mile marathon on a hot, sunny day.

[12] For example, just one clerical error in a multiple regression of one million data points can cause the resulting regression parameter estimates to be considerably different.

[13] That is unless one attempts to hide the lack of information by making judgments that appear to be consistent. One is reminded of Ralph Waldo Emerson's saying, "Foolish consistency is the hobgoblin of small minds".

Lack of Concentration

Another cause of inconsistency is lack of concentration during the judgment process. This can happen if the people making judgments become fatigued[14] or are not really interested in the decision.

Real World Is Not Always Consistent

Still another cause of a high inconsistency ratio is an actual lack of consistency in whatever is being modeled. The real world is rarely perfectly consistent and is sometimes fairly inconsistent. Professional sports is a good example. It is not too uncommon for Team A to defeat Team B, after which Team B defeats Team C, after which Team C defeats Team A! Inconsistencies such as this may be explained as being due to random fluctuations, or to underlying causes (such as match-ups of personnel), or to a combination. Regardless of the reasons, real world inconsistencies do exist and thus will appear in our judgments.

Inadequate Model Structure

A final cause of inconsistency is "inadequate" model structure. Ideally, one would structure a complex decision in a hierarchical fashion such that factors at any level are comparable, within an order of magnitude or so, of other factors at that level. Practical considerations might preclude such a structuring and it is still possible to get meaningful results. Suppose for example, we compared several items that differed by as much as two orders of magnitude. One might erroneously conclude that the AHP scale is incapable of capturing the differences since the scale ranges[15] from 1 to 9. However, because the resulting priorities are based on second, third, and higher order dominances, AHP can produce priorities far beyond an order of magnitude[16]. A higher than usual inconsistency ratio will result because of the extreme judgments necessary. If one recognizes this as the cause,

---

[14] At which point it is time to stop and resume at a later time.
[15] Actually 9.9 using the Expert Choice numerical mode.
[16] For example, if A is nine times B, and B is nine times C, then the second order dominance of A over C is 81 times.

(rather than a clerical error for example), one can accept the inconsistency ratio even though it is greater than 10%.

Necessary but not sufficient

It is important that a low inconsistency not become the goal of the decision-making process. A low inconsistency is necessary but not sufficient for a good decision. It is possible to be perfectly consistent but consistently wrong. It is more important to be accurate than consistent.

Compensatory and Non-Compensatory Decision-making

As mentioned above, AHP is a compensatory decision methodology because alternatives that are deficient with respect to one or more objectives can compensate by their performance with respect to other objectives. Hogarth[17] has categorized ‘decision rules’ for choice into two groups:

(1) strategies that confront the conflicts inherent in the choice situation; and
(2) strategies that avoid the conflicts. Conflict-confronting strategies are compensatory. That is, they allow you to trade off a low value on one dimension against a high value on another. Conflict-avoiding strategies, on the other hand, are non-compensatory. That is, they do not allow trade-offs.

According to Hogarth, the most straightforward, and in many ways most comprehensive strategy (for choice), is the so-called linear compensatory model.

Under a set of not too restrictive assumptions, this (the linear compensatory model) is quite a good choice model from a normative viewpoint. At a descriptive level, the linear model has been shown to be remarkably accurate in predicting individual judgments in both laboratory and applied settings.[18]

---

[17] Robin Hogarth *Judgment and Choice*, John Wiley & Sons, New York, 1987, p. 72.
[18] Ibid., p74.

Up until recently, the linear model as a choice process has been inadequate because, as Hogarth points out, it implies a process of explicit calculations and the trading off of dimensions, which, when there are many alternatives and dimensions, is not feasible for unaided judgment. Even when the number of dimensions and number of alternatives are small, people may still avoid compensatory strategies in making choices. According to Hogarth[19]

> A number of studies have shown, for instance, that preferences based on holistic or intuitive judgment differ from those constructed by use of a linear model; however, the latter are more consistent in the sense that final judgments show less disagreement between different individuals than intuitive evaluations. Intuitive judgment has two sources of inconsistency: in the application of weights attributed to dimensions, and in the aggregation of information across dimensions.

Today, with AHP and readily available computer technology, we can take advantage of the linear compensatory model[20].

Principles and Axioms of the Analytic Hierarchy Process

AHP is built on a solid yet simple theoretical foundation. The basic ‘model’ is one that almost every executive is familiar with — a pie chart. If we draw a pie chart, the whole of the chart represents the goal of the decision problem. The pie is organized into wedges, where each wedge represents an objective contributing to the goal. AHP helps determine the relative importance of each wedge of the pie. Each wedge can then be further decomposed into smaller wedges representing sub-objectives. And so on. Finally, wedges corresponding to the lowest level sub-objectives are broken down into alternative wedges, where each alternative wedge represents how much the alternative contributes to that sub-objective. By adding up the priority for the wedges for the alternatives, we determine how much the alternatives contribute to the organization’s objectives.

---

[19] Ibid., 74.

[20] The AHP is actually more elaborate than a simple linear model because the multiplication of priorities from one level to the next results in what mathematicians refer to as a multilinear model.

AHP is based on three basic principles: decomposition, comparative judgments, and hierarchic composition or synthesis of priorities.[21] The decomposition principle is applied to structure a complex problem into a hierarchy of clusters, sub-clusters, sub-sub clusters and so on. The principle of comparative judgments is applied to construct pairwise comparisons of all combinations of elements in a cluster with respect to the parent of the cluster. These pairwise comparisons are used to derive 'local' priorities of the elements in a cluster with respect to their parent. The principle of hierarchic composition or synthesis is applied to multiply the local priorities of elements in a cluster by the 'global' priority of the parent element, producing global priorities throughout the hierarchy and then adding the global priorities for the lowest level elements (the alternatives).

All theories are based on axioms. The simpler and fewer the axioms, the more general and applicable is the theory. Originally AHP was based on three relatively simple axioms. The first axiom, the reciprocal axiom, requires that, if $P_C(E_A,E_B)$ is a paired comparison of elements A and B with respect to their parent, element C, representing how many times more the element A possesses a property than does element B, then $P_C(E_B,E_A) = 1/ P_C(E_A,E_B)$. For example, if A is 5 times larger than B, then B is one fifth as large as A.

The second, or homogeneity axiom, states that the elements being compared should not differ by too much, else there will tend to be larger errors in judgment. When constructing a hierarchy of objectives, one should attempt to arrange elements in a cluster so that they do not differ by more than an order of magnitude. (The AHP verbal scale ranges from 1 to 9, or about an order of magnitude. The numerical and graphical modes of Expert Choice accommodate almost two orders of magnitude, allowing a relaxation of this axiom. Judgments beyond an order of magnitude generally result in a decrease in accuracy and increase in inconsistency).

The third axiom states that judgments about, or the priorities of, the elements in a hierarchy do not depend on lower level elements. This axiom is required for the principle of hierarchic composition to apply. While the

[21] T. L. Saaty, *Fundamentals of Decision Making and Priority Theory with the Analytic Hierarchy Process*, RWS Publications, Pittsburgh PA., 1994, p 337.

first two axioms are always consonant with real world applications, this axiom requires careful examination, as it is not uncommon for it to be violated. Thus, while the preference for alternatives is almost always dependent on higher level elements, the objectives, the importance of the objectives might or might not be dependent on lower level elements, the alternatives. For example, in choosing a laptop computer, the relative importance of speed vs. weight might depend on the specific alternatives being considered—if the alternatives were about the same weight but differed greatly in speed, then speed might be more important. We say there is feedback from the alternatives to the objectives. There are two basic ways to proceed in those situations where this axiom does *not* apply, that is, when there is feedback. The first involves a formal application of feedback and a supermatrix calculation for synthesis rather than hierarchic composition. This approach is called the Analytic Network Process. For simple feedback (between adjacent levels only), this is equivalent to deriving priorities for the objectives with respect to each alternative, in addition to deriving priorities for the alternatives with respect to each objective. The resulting priorities are processed in a supermatrix, which is equivalent to the convergence of iterative hierarchical compositions. While this approach is extremely powerful and flexible (feedback within levels and between nonadjacent levels can also be accommodated), a simpler approach that usually works well is to make judgments for lower levels of the hierarchy first (or to reconsider judgments at the upper levels after making judgments at the lower level). In so doing, the brain performs the feedback function by considering what was learned at lower levels of the hierarchy when making judgments for upper levels. Thus, an important rule of thumb is to make judgments in a hierarchy from the bottom up, unless one is sure that there is no feedback, or one already has a good understanding of the alternatives and their tradeoffs. Even if this is *not* done, adherence to AHP's fourth axiom (below) as well as the process notion of AHP, can usually lead to appropriate judgments, since an examination of the priorities after a first iteration of the model will highlight those areas where judgments should be revised based on what has been learned.

A fourth axiom, introduced later by Saaty, says that individuals who have reasons for their beliefs should make sure that their ideas are adequately represented for the outcome to match these expectations. While this axiom might sound a bit vague, it is very important because the generality of AHP makes it possible to apply AHP in a variety of ways and adherence to this axiom prevents applying AHP in inappropriate ways. We will illustrate this a bit later.

The simplicity and generality of AHP fit nicely with Ockham's razor, which contends that the simplest of two or more competing theories is preferable. Not only do we contend that AHP's axioms are simpler and more realistic than other decision theories, but that the ratio scale measures that it produces makes AHP more powerful as well.

Expert Choice

We will illustrate AHP and Expert Choice with a simple site location problem. Assume that we want to determine the best retail site within a geographic area for a small ice cream store catering to young children and families. We have narrowed down the site alternatives to three locations: the first one is a suburban shopping center. The second site is in the main business district area of the city, and the third is a suburban mall location. Details regarding each of these alternative sites are presented below:

Suburban Shopping Center

A vacant store location that was formerly a pizza shop is available for $28/sq. ft. per month in a neighborhood "strip" shopping center at a busy highway intersection. The area is populated with 45,000 (mostly middle income, young family) residents of a community who live in townhouses and single family dwellings. The strip center is constantly busy with retail customers of the major supermarket chain, a drug store, a hardware store, a hair stylist/barber shop, and several other small businesses sharing the location. No ice cream shops are located in the community.

The Mall

This location would cost $75/sq. ft. per month. We would be in the main food area of a major suburban mall with 75 retail shops and three “magnet” stores (Sears and two large department stores). The mall is frequented by teens, young mothers, and families usually on weekend days and weekday nights. There are three ice cream stores at various locations within the mall.

Main Street

For $50/sq. ft. per month we can locate our ice-cream store in the ground level of a large high rise office and retail complex. The shop would be in a moderately out of the way corner of the building. The majority of the people frequenting the building and the surrounding area are young professionals who are in the area Monday through Friday only. There is one ice cream store within a ten-block radius of this location.

The information on the three candidate sites can be used to build a basic model. There can be many variations to this model depending on how you choose to structure the problem. There is no one specific right or wrong way to model any decision problem. A basic approach to modeling the site location problem is outlined next.

Developing a Decision Hierarchy

Step 1. Decompose the Problem

The first step in using AHP and the Expert Choice software is to develop a hierarchy by breaking the problem down into its components. The three major levels of the hierarchy (shown in alternative views in Figure 2 and Figure 3) are the goal, objectives, and alternatives.

Goal

The goal is a statement of the overall objective. In our example, to Select the Best Retail Site

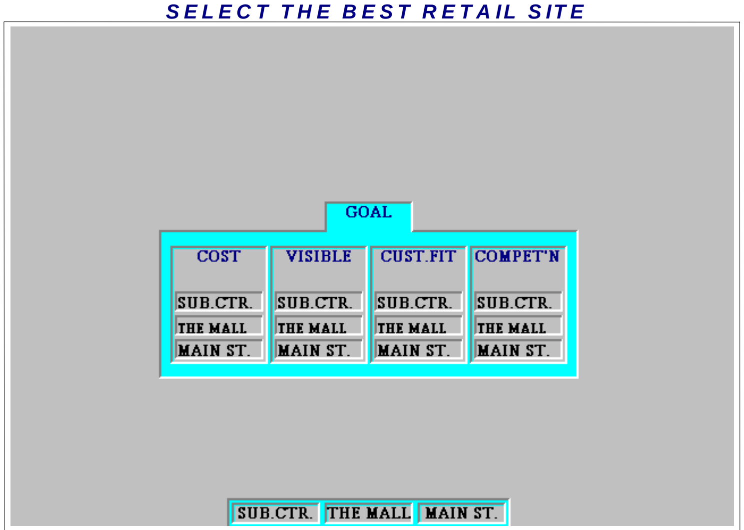
Objectives

What are we trying to achieve in selecting the site? In our example: Cost(low), Visibility, Customer Fit, Competition (lack of).

Alternatives

We consider the feasible alternatives that are available to reach the ultimate goal. In our example the alternatives that have been identified are a Suburban Shopping Center, The Mall, and Main Street.

SELECT THE BEST RETAIL SITE



Abbreviation	Definition
COMPET'N	COMPETITION--# OF COMPETITIVE STORES IN SAME TRADING AREA
COST	COST PER SOUARE FOOT OF RETAIL SPACE
CUST.FIT	CUSTOMER FIT--SITE'S CUSTOMER TRAFFIC VS. TARGET MARKET SPEC'S
MAIN ST.	MAIN STREET--CENTER CITY. OFFICE & RETAIL COMPLEX SITE
SUB.CTR.	SUBURBAN STRIP SHOPPING CENTER
THE MALL	SUBURBAN SHOPPING MALL SITE
VISIBLE	VISIBILITY OF STORE FRONT

Figure 2 – Basic EC Model with Goal, Objectives and Alternatives

More Complex Hierarchies

Expert Choice can easily support more complex hierarchies containing sub-objectives, scenarios or uncertainties, and players. Another variation, the ratings approach, can be used to evaluate a large number of alternatives.

Sub-objectives—

This allows more specificity in the model. By adding sub-objectives you can detail your objectives. Figure 4 shows the model with sub-objectives added for COST, CUSTOMER FIT and COMPETITION.

SELECT THE BEST RETAIL SITE

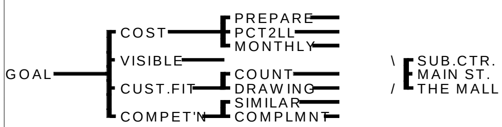
Abbreviation	Definition
COMPET'N	COMPETITION--# OF COMPETITIVE STORES IN SAME TRADING AREA
COST	COST PER SQUARE FOOT OF RETAIL SPACE
CUST.FIT	CUSTOMER FIT--SITE'S CUSTOMER TRAFFIC VS. TARGET MARKET SPEC'S
MAIN ST.	MAIN STREET--CENTER CITY, OFFICE & RETAIL COMPLEX SITE
SUB.CTR.	SUBURBAN STRIP SHOPPING CENTER
THE MALL	SUBURBAN SHOPPING MALL SITE
VISIBLE	VISIBILITY OF STORE FRONT

Figure 3 – Basic EC Model: An Alternative View

Scenarios or Uncertainties—

The importance of different objectives and alternatives may depend on the specific future conditions, which are often difficult to predict. Scenarios can be modeled with Expert Choice allowing you to consider decision alternatives under a variety of circumstances. Scenarios representing the three possible states of the economy, Gloomy Economy, Boom Economy, and Status Quo are shown in Figure 5.

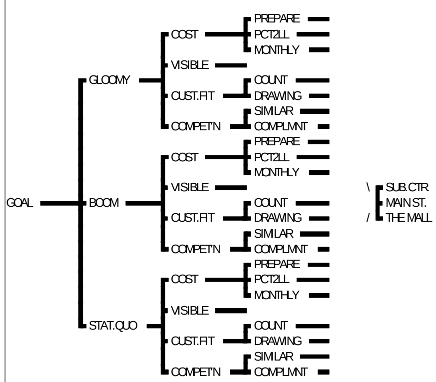
SELECT BEST RETAIL SITE



Abbreviation	Definition
COMPET'N	COMPETITION
COMPLMNT	COMPLIMENTARY STORES THAT WILL ATTRACT CUSTOMERS TO AREA
COST	COST
COUNT	COUNT OF POTENTIAL CUSTOMERS PASSING BY
CUST.FIT	CUSTOMER FIT
DRAWING	ABILITY TO DRAW POTENTIAL CUSTOMERS INTO STORE
MAIN ST.	MAIN ST.
MONTHLY	MONTHLY LEASE COST
PCT2LL	PERCENTAGE OF GROSS MONTNLY INCOME TO GO TO LANDLORD
PREPARE	COST TO PREPARE SITE (INITIAL COST)
SIMILAR	COMPETITION FROM SIMILAR STORES
SUB.CTR.	SUBURBAN CENTER
THE MALL	THE MALL
VISIBLE	VISIBILITY

Figure 4 – Model with Sub-objectives

SELECT BEST RETAIL SITE

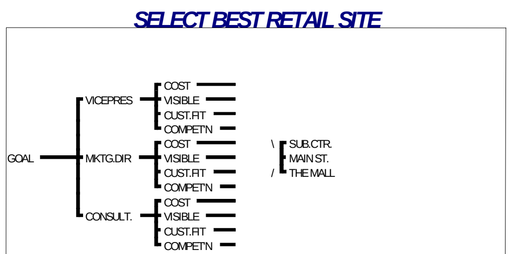


Abbreviation	Definition
BOOM	BOOM ECONOMIC SCENARIO
COMPET'N	COMPETITION
COMPET'N	COMPET'N
COMPLMNT	COMPLIMENTARY STORES THAT WILL ATTRACT CUSTOMERS TO AREA
COST	COST
COST	COST
COUNT	COUNT OF POTENTIAL CUSTOMERS PASSING BY
CUST.FIT	CUSTOMER FIT
CUST.FIT	CUST.FIT
DRAWING	ABILITY TO DRAW POTENTIAL CUSTOMERS INTO STORE
GLOOMY	GLOOMY ECONOMIC OUTLOOK
MAIN ST.	MAIN ST.
MONTHLY	MONTHLY LEASE COST
PCT2LL	PERCENTAGE OF GROSS MONTNLY INCOME TO GO TO LANDLORD

Figure 5 – Model with Scenarios

Players—

Decisions are often made through group consensus, yet it is often difficult for all members of a group to meet, or for each member's opinions to be heard during a meeting. By including a level for players in an EC model, each member's views can be incorporated into the decision-making process. Figure 6 illustrates players that include a Vice-President, Marketing Director, and Consultant



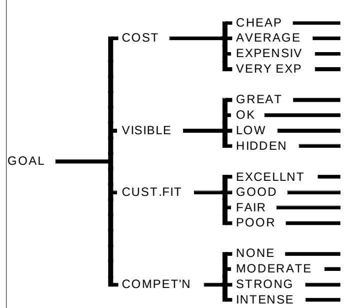
Abbreviation	Definition
COMPET'N	COMPET'N
CONSULT.	SITE LOCATION CONSULTANT
COST	COST
CUST.FIT	CUST.FIT
MAIN ST.	MAIN ST.
MKTG.DIR	MARKETING DIRECTOR
SUB.CTR.	SUB.CTR.
THE MALL	THE MALL
VICEPRES	VICE PRESIDENT OF STORE OPERATIONS
VISIBLE	VISIBLE

Figure 6 – EC Model with Company Players

Ratings Approach for A Large Number of Alternatives—

Some decisions inherently involve a large number of alternatives, which need to be considered. When this is true, the Expert Choice ratings approach easily accommodates a large number of alternatives, such as dozens (or even thousands) of potential sites to compare in a large metropolitan area. Figure 7 and Figure 8 illustrate the ratings approach.

SELECT THE BEST RETAIL SITE



Abbreviation	Definition
EXPENSIV	Expensive
COMPET'N	COMPETITION--# OF COMPETITIVE STORES IN SAME TRADING AREA
COST	COST PER SQUARE FOOT OF RETAIL SPACE
CUST.FIT	CUSTOMER FIT--SITE'S CUSTOMER TRAFFIC VS. TARGET MARKET SPEC'S
GREAT	Great visiblity
HIDDEN	Hidden visibility
LOW	Low visibility
MAIN ST.	MAIN STREET--CENTER CITY, OFFICE & RETAIL COMPLEX SITE
OK	OK visibility
SUB.CTR.	SUBURBAN STRIP SHOPPING CENTER
THE MALL	SUBURBAN SHOPPING MALL SITE
VERY EXP	Very expensive
VISIBLE	VISIBILITY OF STORE FRONT

Figure 7 – EC Model with Large Number of Alternatives

The ratings approach consists of defining "intensities" of achievement or preference with respect to each of the objectives. These intensities are used in place of alternatives in the first stage of the evaluation. For example, instead of comparing the relative preference for two specific alternatives with respect to VISIBILITY, we would compare the relative preference for a non-specific alternative that possesses GREAT visibility to some other alternative that has LOW visibility. This results in measures of preference for the intensities. A ratings "spreadsheet" is then used to evaluate each alternative as to its intensity on each objective. With the ratings approach, pairwise comparisons are made for the objectives, as well as for the intensities under each objective. The results are ratio scale priorities for the importance of each objective, as well as ratio scale priorities for the intensities below each objective. Then, using the ratings "spreadsheet", each alternative is evaluated as to its intensity for each objective. The ratio scale priorities are then summed to give an overall ratio scale measure of the preference for the alternatives.

	COST	VISIBLE	CUST.FIT	COMPET'N	
	.	.	.	.	
	.	.	.	.	
	.	.	.	.	
/INTENSITIES/	0.5088	0.2427	0.1550	0.0935	Total
SUBURBAN CNTR#1	CHEAP	OK	GOOD	STRONG	0.7733
SUBURBAN CNTR#2	CHEAP	LOW	GOOD	MODERATE	0.7192
OLD TOWN AREA	VERY EXP	GREAT	EXCELLNT	INTENSE	0.4581
THE MALL	VERY EXP	GREAT	EXCELLNT	STRONG	0.4775
MAINST/HI RISE	EXPENSIV	OK	GOOD	STRONG	0.4341
NEAR APT. CLUSTER	AVERAGE	LOW	GOOD	MODERATE	0.5041
OFF INTERSTATE	EXPENSIV	OK	EXCELLNT	MODERATE	0.5069
SUBURBAN CNTR#3	AVERAGE	HIDDEN	FAIR	MODERATE	0.4493

Figure 8 – Ratings for a Large Number of Alternatives

Establishing Priorities

After arranging the problem in a hierarchical fashion, the next step is to establish priorities. Each node is evaluated against each of its peers in relation to its parent node; these evaluations are called pairwise comparisons. Referring back to the our basic site selection model in Figure 2:

SELECTING THE BEST RETAIL SITE is the ***parent*** node of
COST, VISIBILITY, CUSTOMER FIT, AND COMPETITION.
COST is a ***parent*** to
MAIN STREET, THE MALL, and SUBURBAN CENTER.

COST, VISIBILITY, CUSTOMER FIT, and COMPETITION are ***peers***.

MAIN STREET, THE MALL, and SUBURBAN CENTER are ***peers***.

Pairwise Comparisons

Pairwise comparisons of the elements at each level of an EC model are made in terms of either:

- Importance—when comparing objectives or players with respect to their relative importance.
- Preference—when comparing the preference for alternatives with respect to an objective.
- Likelihood—when comparing uncertain events or scenarios with respect to the probability of their occurrence.

Pairwise comparisons are basic to the AHP methodology. When comparing a pair of "factors"[22], a ratio of relative importance, preference or likelihood of the factors can be established. This ratio need not be based on some standard scale such as feet or meters but merely represents the relationship of the two "factors" being compared. For example, when looking at two lights, we can judge (without any scientific measurement) that one light is

[22] Factors may be objectives, sub-objectives, scenarios, players, or alternatives.

brighter, or perhaps twice as bright as the other. This may be a subjective judgment, but the two lights can be compared as such.

Most individuals would question the accuracy of any judgment made without using a standard scale. Yet, it has been verified that a number of these pairwise comparisons taken together form a sort of average, the results of which are very accurate. This "average" is calculated through a complex mathematical process using eigenvalues and eigenvectors. The results of this method have been tested experimentally and have been found to be extremely accurate. This method is used in AHP and Expert Choice allowing one to use both subjective and objective data in making pairwise comparisons.

Eigenvalues and Eigenvectors

A little of the Math—Why AHP Uses Eigenvalues and Eigenvectors

Suppose we already knew the relative weights of a set of physical objects like n rocks. We can express them in a pairwise comparison matrix as follows:

$$\underline{A} = \begin{bmatrix} w_1/w_1 & w_1/w_2 & w_1/w_3 & \dots & w_{1/}w_n \\ w_2/w_1 & w_2/w_2 & w_2/w_3 & \dots & w_2/w_n \\ w_3/w_1 & w_3/w_2 & w_3/w_3 & \dots & w_3/w_n \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ w_n/w_1 & w_n/w_2 & w_n/w_3 & \dots & w_n/w_n \end{bmatrix}$$

If we wanted to "recover" or find the vector of weights, [w_1, w_2, w_3, ... w_n] given these ratios, we can take the matrix product of the matrix A with the vector w to obtain[23]:

[23] The matrix product is formed by multiplying, element by element, each row of the first factor, $\underline{A}$, by corresponding elements of the second factor, $\underline{w}$, and adding. Thus, the first element of the product would be:

$$\begin{bmatrix} w_1/w_1 & w_1/w_2 & w_1/w_3 & \dots & w_{1/}w_n \\ w_2/w_1 & w_2/w_2 & w_2/w_3 & \dots & w_2/w_n \\ w_3/w_1 & w_3/w_2 & w_3/w_3 & \dots & w_3/w_n \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ w_n/w_1 & w_n/w_2 & w_n/w_3 & \dots & w_n/w_n \end{bmatrix} * \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ \dots \\ \dots \\ \dots \\ w_n \end{bmatrix} = \begin{bmatrix} nw_1 \\ nw_2 \\ nw_3 \\ \dots \\ \dots \\ \dots \\ nw_n \end{bmatrix}$$

$$\underline{A} \quad * \ \underline{w} \ = \ n\,\underline{w}$$

If we knew A, but not w, we could solve the above for w. The problem of solving for a nonzero solution to this set of equations is very common in engineering and physics and is known as an eigenvalue problem:

$$\underline{\mathbf{A}}\ \underline{\mathbf{w}} = \lambda\,\underline{\mathbf{w}}$$

The solution to this set of equations is, in general found by solving an nth order equation for l. Thus, in general, there can be up to n unique values for l, with an associated w vector for each of the n values.

In this case however, the matrix A has a special form since each row is a constant multiple of the first row. For such a matrix, the rank of the matrix is one, and all the eigenvalues of A are zero, except one. Since the sum of the eigenvalues of a positive matrix is equal to the trace of the matrix (the sum of the diagonal elements), the non zero eigenvalue has a value of n, the size of the matrix. This eigenvalue is referred to as $\boldsymbol{\lambda_{max}}$.

Notice that each column of $\underline{A}$ is a constant multiple of $\underline{w}$. Thus, $\underline{w}$ can be found by normalizing any column of $\underline{A}$.

The matrix $\underline{A}$ is said to be strongly consistent in that

$a_{ik}a_{kj} = a_{ij}$ for all i,j.

$(w_1/w_1)*w_1 + (w_1/w_2)*w_2 + ...+ (w_1/w_n)*w_n = nw_1$. Similarly, the second element would be $(w_2/w_1)*w_1 + (w_2/w_2)*w_2 + ...+ (w_2/w_n)*w_n = nw_2$. The n^{th} element would be nw_n. Thus, the resulting vector would be n$\underline{w}$.

Now let us consider the case where we do *not* know $\underline{w}$, and where we have only estimates of the $a_{ij's}$ in the matrix $\underline{A}$ and the strong consistency property most likely does not hold. (This allows for small errors and inconsistencies in judgments). It has been shown that for any matrix, small perturbations in the entries imply similar perturbations in the eigenvalues, thus the eigenvalue problem for the inconsistent case is:

$\underline{A}\ \underline{w} = \lambda_{max}\ \underline{w}$,

where $\boldsymbol{\lambda_{max}}$ will be close to n (actually greater than or equal to n) and the other λ's will be close to zero. The estimates of the weights for the activities can be found by normalizing the eigenvector corresponding to the largest eigenvalue in the above matrix equation.

The closer λ_{max} is to n, the more consistent the judgments. Thus, the difference, $\lambda_{max} - n$, can be used as a measure of inconsistency (this difference will be zero for perfect consistency). Instead of using this difference directly, Saaty defined a consistency index as:

$(\lambda_{max} - n)/(n-1)$

since it represents the average of the remaining eigenvalues.

In order to derive a meaningful interpretation of either the difference or the consistency index, Saaty simulated random pairwise comparisons for different size matrices, calculating the consistency indices, and arriving at an average consistency index for random judgments for each size matrix. He then defined the consistency ratio as the ratio of the consistency index for a particular set of judgments, to the average consistency index for random comparisons for a matrix of the same size. Forman[24] performed additional simulations and calculated indices for cases with missing judgments.

Since a set of perfectly consistent judgments produces a consistency index of 0, the consistency ratio will also be zero. A consistency ratio of 1 indicates consistency akin to that, which would be achieved if judgments were not made intelligently, but rather at random. This ratio is called the <u>in</u>consistency ratio in Expert Choice since the larger the value, the more inconsistent the judgments.

---

[24] E. H. Forman, "Random Indices for Incomplete Pairwise Comparison Matrices" *European Journal of Operations Research* Vol. 48, #1, 1990, pp. 153-155

Note: Other methods to estimate activity weights, such as least squares and log least squares have been suggested. While these methods produce results that are similar to the eigenvector approach, no other method maintains the reciprocal property of the pairwise comparison matrix (known as weak consistency), nor produces a comparable measure of inconsistency.

Note: An approximation to the Eigenvector method suitable for hand calculations is available (for example, Dyer and Forman[25]). While this approximation is reasonable when the judgments are relatively consistent, it may not be so for inconsistent judgments and is thus not recommended unless a computer and software are not available.

Because of the reciprocal property of the comparison matrix, the eigenvector problem can be solved by raising the matrix to the n^{th} power, and taking the limit as n approaches infinity. The matrix will always converge. Saaty has shown that this corresponds to the concept of dominance walks. The dominance of each alternative along all walks of length k, as k goes to infinity, is given by the solution to the eigenvalue problem[26].

Three pairwise comparison modes

Expert Choice allows you to enter judgments in either numerical, graphical, or verbal modes. Each judgment expresses the ratio of one element compared to another element. When making comparisons in a social, psychological, or political context, you may wish to use the verbal comparison mode. Verbal judgments are easier to make, and for qualitative or value driven comparisons, easier to justify. When comparing economic or other measurable factors, the numerical or graphical comparison modes may be preferred, although it is perfectly acceptable to use the verbal mode in that case as well[27].

---

[25] Dyer, Robert F. and Forman, Ernest H., *An Analytic Approach to Marketing Decisions,* Prentice Hall 1991,pp. 92-93.

[26] For further information, see Saaty pp. 78-121.

[27] The verbal mode is not as accurate and typically requires more judgments than the numerical or graphical mode in order to improve accuracy.

Numerical judgments

When comparing properties that lend themselves naturally to a numerical scale, one can use the numerical mode to enter the judgments. In the numerical scale, 1.0 implies that the elements are equally important, 2.0 that one element is twice as important as the other, and 9.0 that one element is nine times as important as the other. These are absolute numbers that tell us, for example, which of two stones is the heavier and how much heavier it is. Thus, a numerical judgment of 5.0 says that the first stone is five times heavier than the second. If the disparity between elements in a group is so great that they are not of the same "order of magnitude" that is, some elements in the group are more than 9.0 times greater than some other elements in the group, they should be put into clusters of like elements. Alternatively, Expert Choice allows expansion of the numerical scale to a ratio of 99.9 to 1; however people are not as accurate in making judgments when the elements differ by ratios of 10 to 1 or more.

Graphical judgments

The graphical pairwise comparison scale can be used to express the relationships between two elements as the ratio of the lengths of two bars. Judgments are entered in the graphical mode by dragging and adjusting the relative lengths of the two bars (each representing one of the factors in the pairwise comparison). A pie chart changes to reflect the same relative proportion as you drag the bars (see Figure 9).

Verbal judgments

The nine point verbal scale used in Expert Choice is presented in Table 1.

Table 1 – EC Pairwise Comparison Scale.

Numerical Value	Verbal Scale	Explanation
1.0	Equal importance of both elements	Two elements contribute equally
3.0	Moderate importance of one element over another	Experience and judgment favor one element over another
5.0	Strong importance of one element over another	An element is strongly favored
7.0	Very strong importance of one element over another	An element is very strongly dominant
9.0	Extreme importance of one element over another	An element is favored by at least an order of magnitude
2.0, 4.0, 6.0, 8.0	Intermediate values	Used to compromise between two judgments

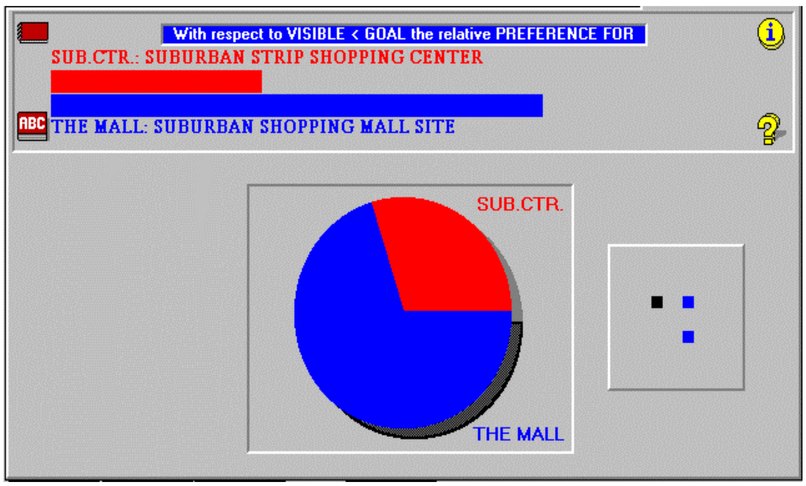


Figure 9 – Pairwise Graphical Comparison

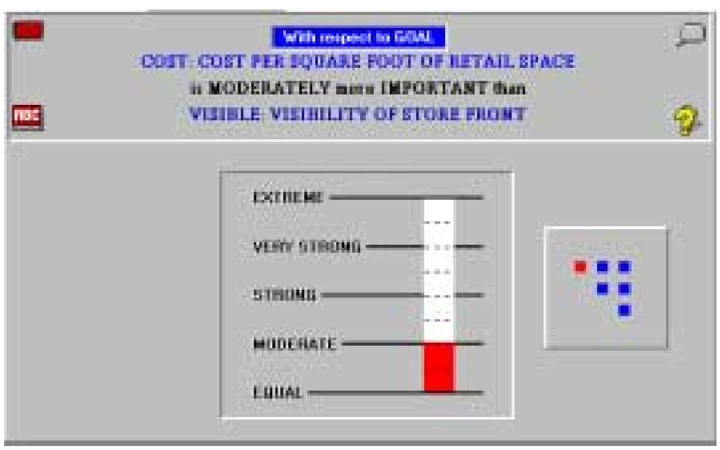


Figure 10 – Verbal Judgment

A verbal judgment in Expert Choice is shown in Figure 10.

Whereas numerical and graphical judgments are in and of themselves ratios and hence possess the ratio level of measurement, the same is not true for the verbal scale. The verbal scale is essentially an ordinal scale. When a decision maker judges A to be Strongly more important than B we know that A is 'more' important than B, but we do not know the interval between A and B or the Ratio of A to B. Studies have shown that relative (pairwise) *verbal* judgments can produce accurate, ratio scale priorities from what are basically imprecise, ordinal judgments, provided that redundant judgments are included in the calculations. Redundancy helps to reduce the average effect of errors in a manner analogous to the way that taking the average of a sample of measurements will produce an estimate of the mean that is likely to be closer to the true mean than only one judgment (i.e., no redundancy.) In addition to reducing the effect of the usual type of errors in measurement, this procedure also reduces the effect of the fuzzy nature of the ordinal scale and different interpretations of the scale by different decision-makers.

While relative pairwise judgments can be made numerically or graphically, verbal judgments are important in decision-making because humans have learned to use and are comfortable in using words to measure the intensity of feelings and understanding with respect to the presence of a property. For example, one might be more comfortable in saying that one fruit tastes moderately sweeter than another fruit than in saying that one fruit tastes three times sweeter than another fruit. Because complex crucial decision situations often involve too many dimensions for humans to synthesize intuitively, we need a way to synthesize over the many dimensions of such decisions. The derivation of ratio scale priorities from verbal judgments makes this possible.

How do we know that this method for deriving the priorities is accurate?

First because it is based on a sound mathematical foundation discussed above, and second by numerous validation experiments. Saaty performed many such experiments, with members of his family, visitors to his house, colleagues at work, and people attending seminars. It was standard practice for a visitor to the Saaty household to be asked to pairwise compare the weights of different size rocks or different suitcases. In one experiment Saaty placed chairs at various distances (9, 15, 21, and 28 yards) from a light source in his back yard to see if his wife and young children, standing at the light source, could judge the relative brightness of the chairs.[28] The results achieved with pairwise verbal judgments (see Table 2) were in very close agreement with the inverse square law of optics, which says that relative brightness is inversely proportional to the square of the distance of the object from the light source.

Table 2 – Results of Brightness of light on Chairs Experiment

Chair	Estimates from Wife's Judgments	Estimates from Sons' Judgments	Results from applying Inverse Square Law of Physics
1	0.61	0.62	0.61
2	0.24	0.22	0.22
3	0.10	0.10	0.11
4	0.05	0.06	0.06

Wedley has performed validation studies estimating the relative color intensities of objects. An Area Validation Study has shown not only is it possible to derive fairly accurate ratio scale priorities from verbal judgments, but that the redundancy in the pairwise process increases accuracy significantly. Consider the following analogy. Suppose you were allocating funds for environmental quality purposes and wanted to determine the relative funding for clean air, clean water, noise reduction, industrial dumps, and acid rain. As the analogy, suppose your insight about the relative needs coincide with the areas of the five objects in Figure 11.

[28] Saaty *The Analytic Hierarchy Process*, p. 39.

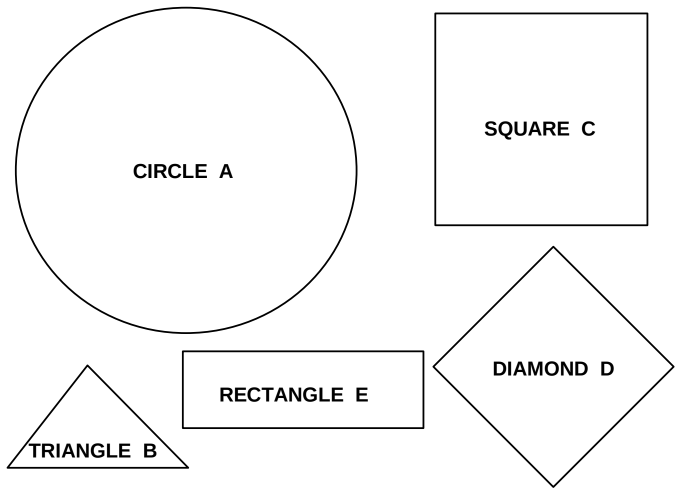


Figure 11 – Estimating Relative Areas Using Words

In this analogy you could look at these objects and estimate their relative sizes numerically, but the analogy is meant to show how words, instead of numbers, can be used to derive accurate priorities for qualitative factors. Using the words, moderate, strong, very strong, and extreme[29], many individuals and groups have made judgments about the relative sizes of these five objects. For example, the verbal judgments made by a group at one organization are shown in Figure 11.

A was judged to be Very Strongly larger than B (represented by a 7), Moderately larger than C (represented by a 3), and so on. The matrix of judgments corresponding to these verbal judgments and used to find the principle right hand eigenvector is shown in Figure 13.

[29] In pairwise relative comparisons, with redundancy.

Area Validation -- March 21, 1991

Node: 0

Compare the relative IMPORTANCE with respect to: GOAL

1=EQUAL 3=MODERATE 5=STRONG 7=VERY STRONG 9=EXTREME

1	A	9	8	(7)	6	5	4	3	2	**1**	2	3	4	5	6	7	8	9	B
2	A	9	8	7	6	5	4	(3)	2	**1**	2	3	4	5	6	7	8	9	C
3	A	9	8	7	6	5	(4)	3	2	**1**	2	3	4	5	6	7	8	9	D
4	A	9	8	7	(6)	5	4	3	2	**1**	2	3	4	5	6	7	8	9	E
5	B	9	8	7	6	5	4	3	2	**1**	2	3	(4)	5	6	7	8	9	C
6	B	9	8	7	6	5	4	3	2	**1**	2	3	(4)	5	6	7	8	9	D
7	B	9	8	7	6	5	4	3	2	**1**	2	(3)	4	5	6	7	8	9	E
8	C	9	8	7	6	5	4	3	(2)	**1**	2	3	4	5	6	7	8	9	D
9	C	9	8	7	6	5	(4)	3	2	**1**	2	3	4	5	6	7	8	9	E
10	D	9	8	7	6	5	4	(3)	2	**1**	2	3	4	5	6	7	8	9	E

Abbreviation	Definition
Goal	Area Validation -- March 21, 1991
A	Circle
B	Triangle
C	Circle
D	Diamond
E	Rectangle

Figure 12 – Area Validation

However, the numeric values in the matrix can *not* be assumed to be accurate since they represent imprecise verbal judgments (really only an ordinal level of measure).

Algebraically, a minimum of n-1 or 4 comparisons is required to derive the relative areas. Any four judgments that "span" each of the elements being compared are sufficient. For example, *if* one assumes that the numerical representations of the verbal judgments shown in the first row of the matrix are accurate, the resulting priorities would be[30] A = 52.8%; B = 7.5%; C = 17.6%; D = 13.2%; and E = 8.8%. Priorities derived as the normalized values of the eigenvector corresponding to the normalized largest eigenvalue of the full

[30] Calculated by solving simultaneous equations: A = 7B, A=3C, A=4D, A=5E, A+B+C+D+E = 1.

$$A = \begin{bmatrix} 1 & 7 & 3 & 4 & 6 \\ 1/7 & 1 & 1/4 & 1/4 & 1/3 \\ 1/3 & 4 & 1 & 2 & 4 \\ 1/4 & 4 & 1/2 & 1 & 3 \\ 1/6 & 3 & 1/4 & 1/3 & 1 \end{bmatrix}$$

Figure 13 – Numerical Representation of Verbal Judgments

pairwise comparison matrix are almost always more accurate than results based on the minimal number of required judgments. This is indeed the case for the set of judgments shown in Figure 13, as can be seen in Table 3 by comparing the priorities derived from the minimal set of judgments (without redundancy) to the actual priorities and then comparing the priorities derived from the full set of judgments to the actual priorities.

Table 3 – Comparisons of Estimated vs. Actual Priorities

FIGURE	Estimates from Verbal Judgments W/O Redundancy	Estimates from Verbal Judgments With Redundancy	Actual
A	52.8	49.4	47.5
B	7.5	4.5	4.9
C	17.6	22.9	23.2
D	13.2	15.4	15.1
E	8.8	7.8	9.3

The sum of squares of the error in the priorities derived from the full set of judgments (those shown in) is 6.2 while the sum of squares of the error in the priorities derived from just the top row of judgments is 70.07.

Notice the priorities derived from the full set of judgments do *not* necessarily agree with the numerical representation of any one judgment. For example, the ratio of the derived priorities of Circle-A and Triangle-B is about 10.9, which is closer to the true ratio of 9.7 than is the 7.0 numerical representation of the verbal judgment VERY STRONG.

We will conclude our discussion of priority derivation with two observations. Priorities derived for the five areas in this validation example from numerical or graphical judgments should be even closer since this is an objective problem with known answers. However, verbal judgments are often more appropriate when judging qualitative factors and all crucial decisions have qualitative factors that must be evaluated. Secondly, we have noted that both numerical and graphical judgments are in and of themselves ratios and hence possess the ratio level of measurement, the verbal scale is essentially an ordinal scale, which *can* be used to produce accurate ratio scale priorities. Do not, however, take this for granted! There may be cases where intervals or ratios of the priorities resulting from verbal judgments do not adequately represent the decision maker(s) feelings. It is incumbent upon the decision maker(s) to examine the resulting priorities and if they do not adequately represent the decision maker(s) feelings, to revise the judgments in either the graphical or numerical modes, preferably the former.

Preference of Alternatives with respect to Objectives

We usually evaluate the preference for the alternatives with respect to the objectives before evaluating the importance of the objectives. This 'bottom up' approach is recommended so that we get a better understanding of the alternatives just in case our judgments about the importance of the objectives are dependent on the alternatives (see discussion of AHP's third axiom on page 51). In our example, we would determine our preferences for MAIN STREET, THE MALL, and SUBURBAN CENTER with respect each of the four objectives. Considering COST, we might proceed as follows. Since cost is an 'objective' objective, we can refer to financial data. Monthly rent on MAIN STREET is $50 per square foot, while rent in THE MALL is $75 per square foot, and the SUBURBAN CENTER is $28 per square foot. While we could enter this data directly, we can also factor in

our subjective interpretation of these costs, reflecting a 'non-linear utility function'. Even though THE MALL is about 2.6 times more costly than the

Figure 14 – Preference for Alternatives with Respect to Cost

SUBURBAN CENTER (based on the actual cost data), we might judge that our *preference* for a cost of $28 per square foot per month is perhaps six times more preferable than $50 per square foot per month. Similarly, even though the SUBURBAN CENTER is 1.8 times more preferable than MAIN STREET based on a linear interpretation of the actual data, we might judge that a rent of $28 per square foot is three times more preferable than a rent of $50 per square foot. Finally, we might judge MAIN STREET to be about 1.5 times more preferable than THE MALL, a judgment that corresponds to a linear interpolation of the ratio of the respective costs. These judgments and resulting priorities are shown in Figure 14.

After judgments about the preferences for the alternatives have been made with respect to the COST objective, we derive priorities for the alternative sites with respect to each of the remaining objectives using either numerical, graphical, or verbal judgments pairwise judgments, or actual data.

Importance of the Objectives with Respect to the Goal

Verbal judgments about the relative importance of each objective are shown in Figure 15. In this example, we judged COST to be moderately more important to us than VISIBILITY with respect to their parent node, the GOAL of CHOOSING THE BEST RETAIL SITE. In other words, it is moderately more important to us to have an affordable location than one that is highly visible. This judgment can be based on our intuition—we know that people will find our shop, due to other factors (promotion, word of mouth, and so on) even if the storefront lacks high visibility; or we can base our judgment on objective data. Our financial analysis, which includes the rent of the sites, makes it clear that COST is more important since we would

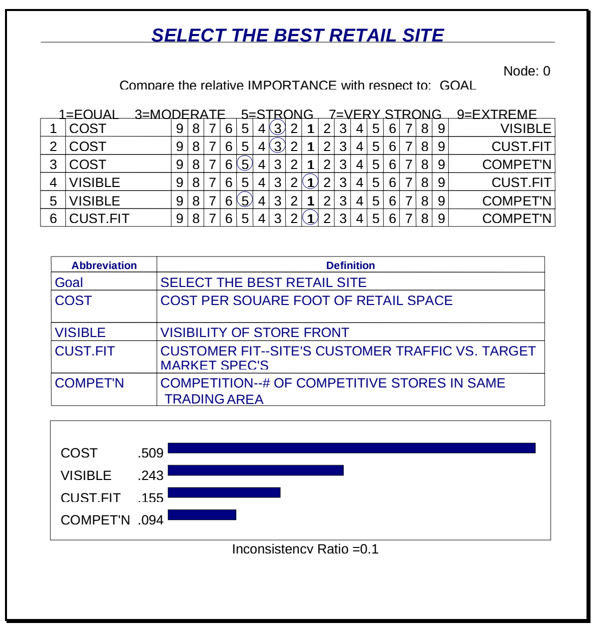


Figure 15 – Judgments about Importance of Objectives and Resulting Priorities.

be in financial difficulty if we were to choose the high visibility location which is very costly. Each objective is evaluated with respect to every other objective in a similar fashion, using relevant subjective or objective judgments. The result of these judgments is a prioritization (or weighting) of the objectives as shown in Figure 15.

Synthesis

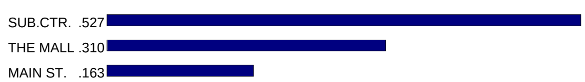
Once judgments have been entered for each part of the model, the information is synthesized to achieve an overall preference. The synthesis produces a report, which ranks the alternatives in relation to the overall goal. This report includes a detailed ranking showing how each alternative was evaluated with respect to each objective. Figure 16 shows the details followed by a ranking of the alternatives.

For our example, the Synthesis shows SUBURBAN CENTER to be the BEST RETAIL SITE. We can examine the details of this decision to see that this site alternative was chosen because it offered better customer fit at a lower cost and with less competition, thus favorably satisfying three of the objectives. Although THE MALL location provided better visibility, it was very expensive and had heavier competition. Expert Choice has helped us determine that the better visibility was not worth the added cost. It is important to note that it would be wrong to conclude that the SUBURBAN CENTER was best overall because it was better on three out of the four objectives. If cost were less important, THE MALL might be best overall as we can see with a sensitivity analysis.

SELECT THE BEST RETAIL SITE

Synthesis of Leaf Nodes with respect to GOAL
Ideal Mode
OVERALL INCONSISTENCY INDEX = 0.07

LEVEL 1	LEVEL 2	LEVEL 3	LEVEL 4	LEVEL 5
COST =.509				
	SUB.CTR.=.509			
	THE MALL=.093			
	MAIN ST.=.154			
VISIBLE =.243				
	SUB.CTR.=.104			
	THE MALL=.243			
	MAIN ST.=.027			
CUST.FIT=.155				
	SUB.CTR.=.155			
	THE MALL=.155			
	MAIN ST.=.031			
COMPET'N=.094				
	SUB.CTR.=.094			
	THE MALL=.016			
	MAIN ST.=.055			



Abbreviation	Definition
GOAL	
COMPET'N	COMPETITION--# OF COMPETITIVE STORES IN SAME TRADING AREA
COST	COST PER SQUARE FOOT OF RETAIL SPACE
CUST.FIT	CUSTOMER FIT--SITE'S CUSTOMER TRAFFIC VS. TARGET MARKET SPEC'S
MAIN ST.	MAIN STREET--CENTER CITY, OFFICE & RETAIL COMPLEX SITE
SUB.CTR.	SUBURBAN STRIP SHOPPING CENTER

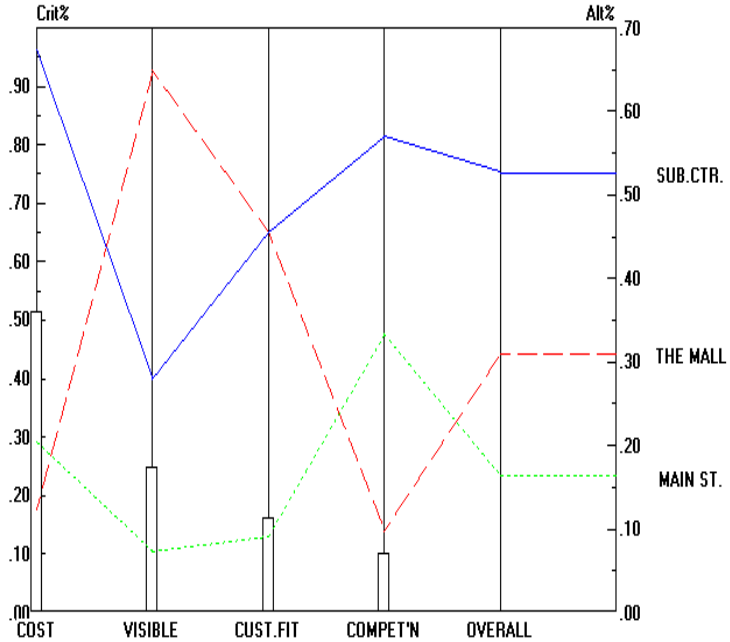
Figure 16 – Synthesis for Site Location Problem[31]

Sensitivity

Sensitivity analyses can be performed to see how well the alternatives performed with respect to each of the objectives as well as how sensitive the alternatives are to changes in the importance of the objectives.

[31] Ideal and Distributive modes of synthesis are discussed on page 151.

Performance Sensitivity w.r.t. GOAL for nodes below GOAL



Abbreviation	Definition
COST	COST PER SQUARE FOOT OF RETAIL SPACE
VISIBLE	VISIBILITY OF STORE FRONT
CUST.FIT	CUSTOMER FIT--SITE'S CUSTOMER TRAFFIC VS. TARGET MARKET SPEC'S
COMPET'N	COMPETITION--# OF COMPETITIVE STORES IN SAME TRADING AREA

SUB.CTR.	SUBURBAN STRIP SHOPPING CENTER
THE MALL	SUBURBAN SHOPPING MALL SITE
MAIN ST.	MAIN STREET--CENTER CITY, OFFICE & RETAIL COMPLEX SITE

Ideal Mode

Figure 17 – Performance Sensitivity

The performance sensitivity shows (see Figure 17) the relative importance of each of the objectives as bars , and the relative preference for each alternative with respect to each objective as the intersection of the alternatives' curves with the vertical line for each objective. The overall alternative preferences are shown at the right. As with all AHP priorities, these priorities are ratio scale priorities meaning that not only do the priorities show order, but differences and ratios are meaningful as well. The SUB.CTR is best on three of the four objectives, and best overall. If however, VISIBILITY were to become more important, then The MALL might become the preferred alternative.

Figure 18 shows a gradient sensitivity analysis of the results with respect to the importance of the VISIBILITY objective. The graph shows that the current priority for VISIBILITY is a little less than .25 (see vertical solid red line). The height of the intersection of this dashed line with the alternative lines shows the alternatives' priorities. Thus, SUB. CTR. is the preferred alternative. If VISIBILITY were to become more important, then your overall preference for SUB.CTR decreases while that of THE MALL increases. If the priority of VISIBILITY were to increase above .53, then THE MALL would be the preferred alternative. However, since it would take a significant change in the priority of VISIBILITY in order to change the ranking of the alternatives, we can say that the results are not very sensitive to small changes in the priority of VISIBILITY. Other types of sensitivity analysis will be discussed later.

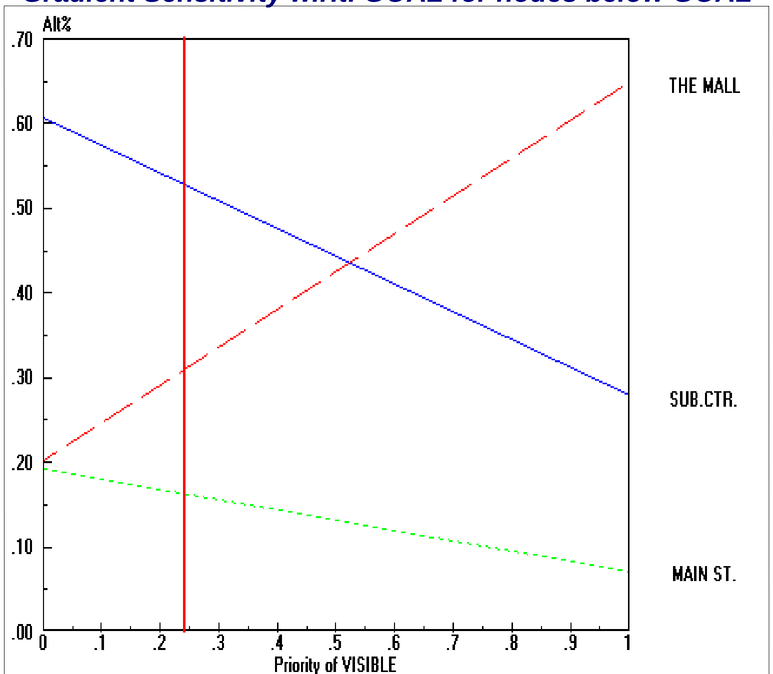
A Typical Application of AHP/EXPERT CHOICE

Choosing a Coast Guard Weapon Patrol Boat Fleet – Background

The United States Coast Guard was operating an aging fleet of small weapon patrol boats (WPBs)[32]. The fleet of WPBs was between twenty and twenty five years old. The original projected life span of these vessels was twenty years; however, several hulls had renovations enabling the useful life to exceed the original projections. Still, the time had come when patching the existing hulls was no longer a reasonable alternative and the fleet had to be replaced. The fleet size was 76 boats but the Coast Guard felt it had obtain at least 90 hulls to provide the necessary level of service required by their operational missions. Regardless of the type of vessel

[32] This study was performed by Edward Hotard and Benjamin Keck, officers in the U.S. Coast Guard assigned to graduate studies at George Washington University. The analysis and conclusions reflect their personal opinions and should in no way be construed as being the opinion of the Commandant of the Coast Guard or the official opinion of any member of the U.S. Coast Guard.

Gradient Sensitivity w.r.t. GOAL for nodes below GOAL



Abbreviation	Definition
COST	COST PER SQUARE FOOT OF RETAIL SPACE
VISIBLE	VISIBILITY OF STORE FRONT
CUST.FIT	CUSTOMER FIT--SITE'S CUSTOMER TRAFFIC VS. TARGET MARKET SPEC'S
COMPET'N	COMPETITION--# OF COMPETITIVE STORES IN SAME TRADING AREA

SUB.CTR.	SUBURBAN STRIP SHOPPING CENTER
THE MALL	SUBURBAN SHOPPING MALL SITE
MAIN ST.	MAIN STREET--CENTER CITY, OFFICE & RETAIL COMPLEX SITE

Ideal Mode

Figure 18 – Gradient Sensitivity Analysis for the Visibility Objective

purchased, each of the replacements would cost at least five million dollars and the projected system life cycle cost was expected to exceed one billion dollars.

The Office of Management and Budget (OMB) requires that major system procurements follow a set of guidelines (entitled OMB A-109) specifying that all reasonable alternatives be investigated before a new or replacement system is selected. The Department of Transportation had designated the Coast Guard WPB acquisition project as a major system acquisition, which came under the purview of OMB A-109. In accordance with OMB A-109 the Coast Guard had determined that the alternatives consisted of several advanced hull forms. The criteria by which the alternative hulls were to be judged were not specified by OMB A-109 and the agency was left with considerable latitude in determining the criteria they wished to use.

Alternatives

The alternative vessel types that were identified for consideration included: a hydrofoil, a small water area twin hull ship (SWATH), a surface effect ship (SES), a planing hull, and a monohull. The present hull form, WPB-95, was included at the request of the Acquisition Project Officer as a baseline in order to observe any anticipated improvements that would result by replacing existing craft with more advanced hull forms.

HYDROFOIL – The hydrofoil was the most technologically advanced of the six candidates. It relied on submerged foils to provide lift to the hull in the same manner that a wing provides lift to an airplane. With less hull surface in contact with the water, the friction or drag of the water against the surface of the hull is reduced allowing for greater speed. Hydrofoils typically have a gas turbine main propulsion plant. The use of a gas turbine requires high speed reduction gears which adds a great deal of complexity to the propulsion drive train. Hydrofoils are very fast, often capable of exceeding 60 knots, an important consideration for use in the war against drug smugglers. Since much of the hull is not in contact with the water the motion of the boat due to wave motions is reduced and a very stable platform results. However, the high speeds require specially designed cavitating screws, more sophisticated control systems, and highly skilled operational and maintenance personnel. The hydrofoil is also one of the more expensive alternatives.

SWATH – A Small Water Area Twin Hull vessel at first glance appears to be similar to a recreational pontoon boat. The pontoons of the SWATH ride not on but beneath the water's surface. The advantage of this type of hull form is that the vessel's center of buoyancy is placed below the surface wave action resulting in a dramatically more stable vessel. The twin hulls however, increase the wetted surface and drag on the hull. Consequently this hull form requires more power per knot than conventional single hull concepts. In addition, the reduced waterplane area produces a vessel that sits deeper in the water and draws a relatively greater draft than the other hulls being considered. While the catamaran style is highly stable over most sea conditions, there are dynamic stability questions that must be resolved when the vessel operates in heavy seas. Unlike single hull designs, its righting arm is comparatively great until it reaches a limiting angle of heel; at this point the righting arm vanishes much more quickly than in other types of hulls and the vessel is subject to being capsized.

SES – The Surface Effect Ship is similar to a hovercraft. The main difference between hovercraft and an SES are the rigid side hulls of the SES. Like the hovercraft, the SES relies on a cushion of air beneath the hull to lift a portion of the hull out of the water, thereby reducing the drag, which results in increased speed. There are, however, several major flaws in this concept. The air under the hull acts as an undampened spring, resulting in a poor ride when sea waves approach the natural frequency of the vessel. In addition, auxiliary motors and fans are required to create the air cushion to lift the vessel out of the water, which adds to the complexity, weight and cost of the ship.

PLANING – The planing hull concept is an evolution of present hull forms that improves dynamic lift and reduces drag. By reducing the resistance of the hull as it is forced through the water the vessel can obtain greater speed and fuel efficiency. The use of aluminum to reduce weight and turbo-charged diesel engines for power, are part of the evolution of ship design that the planing hull has embraced. The main advantages of the planing hull are that it deals with a known technology, has a lesser draft, a high speed in low to moderate seas and moderate cost. The main disadvantages are speed degradation in higher sea states and poor ride quality overall in comparison to some of the more stable advanced designs.

MONOHULL – The monohull is a compromise between what was the existing fleet of weapon patrol boats and the planing hull. The use of aluminum is minimized, the hull is given a sharper rise, and the engines are less powerful than those used in the planing hull. However, the monohull is still evolutionary. The main advantage of the monohull is its cost. Its primary disadvantage is its generally reduced capabilities.

WPB 95 – The current patrol boat configuration was included in the analysis in order to serve as a baseline.

Objectives

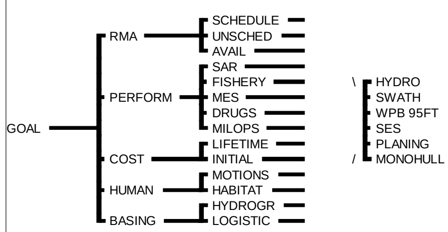
The problem was to determine which of the above vessels "best" met the coast guard's objectives. Based on the authors' knowledge and experience with small patrol craft and the characteristics of the alternatives under consideration, five main objectives were identified:

1. RMA – Reliability, Maintainability, and Availability;
2. Performance;
3. Cost;
4. Human Factors;
5. Basing.

Sub-objectives

Sub-objectives were identified for each of these objectives. These are shown in Figure 19 and will be explained below. The AHP analysis consists of making pairwise comparisons between pairs of factors at one level of the model with respect to the immediately preceding level. Local priorities are derived from each set of pairwise comparisons. Global priorities, relative of the goal, are then calculated and used to determine the ratio scale priorities for the alternatives.

OPTIMAL PATROL CRAFT



Abbreviation	Definition
AVAIL	AVAILABILITY
BASING	HOME PORT CRITERIA
COST	COST
DRUGS	DRUG INTERDICTION
FISHERY	FISHERY ENFORCEMENT
HABITAT	HABITABILITY
HUMAN	HUMAN FACTORS
HYDRO	HYDROFOIL
HYDROGR	HYDROGRAPHICS - WATER DEPTH
INITIAL	ACQUISITION COSTS
LIFETIME	LIFE CYCLE COSTS
LOGISTIC	LOGISTIC
MES	MARINE ENVIRONMENT SAFETY
MILOPS	MILITARY OPERATIONS
MONOHULL	MONOHULL
MOTIONS	SEA STATE MOTIONS
PERFORM	MISSION PERFORMANCE
PLANING	PLANING HULL VESSEL
RMA	RELIABILITY, MAINTAINABILITY, AVAILABILITY
SAR	SEARCH AND RESCUE
SCHEDULE	SCHEDULED MAINTENANCE
SES	SURFACE EFFECT SHIP
SWATH	SMALL WATER AREA TWIN HULL
UNSCHED	UNSCHEDULED MAINTENANCE

Figure 19 – AHP Model for Evaluating Coast Guard Patrol Vessels

Reliability, Maintainability, and Availability

The vessels being considered must be reliable, maintainable, and available (RMA). Three sub-sub-objectives were identified for RMA: scheduled maintenance, unscheduled maintenance, and availability.

Derivation of preferences with respect to lowest level sub-objectives

Relative PREFERENCES for the alternatives relative to the each of these sub-sub-objectives were derived from pairwise verbal comparisons, based on data obtained from computer simulations and empirical evidence.

Pairwise verbal judgments

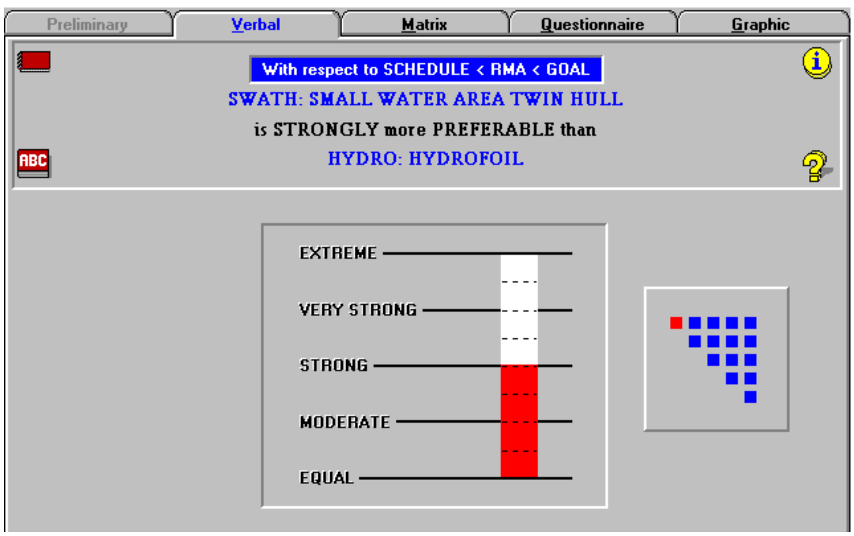


Figure 20 – Pairwise verbal comparison

Figure 20 shows a verbal pairwise judgment in which the small water area twin hull vessel was judged to be strongly more preferable than the hydrofoil with respect to the scheduled maintenance sub-objective. Even when data for comparing the alternatives is available, as is the case for scheduled maintenance, the relative preference is not necessarily linearly related to the data. Another way of saying this, is that the decision maker's "utility curve" for scheduled maintenance is not necessarily linear. Verbal judgments are a natural way of expressing this utility.

Verbal judgments can be represented numerically

Similarly, judgments were made about the relative preference between each pair of alternatives, with respect to the scheduled maintenance sub-objective. These are represented numerically in Figure 21. There is a subtle but important difference between representing verbal judgments numerically and numerical judgments. When representing verbal judgments numerically, a judgment of 'STRONG' is shown as a 5. The decision maker(s) did *not* judge factor A to be five times factor B! *Provided there is suitable redundancy* and that the judgments are reasonably consistent, the eigenvector calculation of priorities based on the full set of pairwise verbal comparisons may result in factor A being only 3 times factor B even though the word 'STRONG' had a numerical representation of 5 (see discussion on page70). Were numerical judgments to be made under similar circumstances, a judgment that factor A is 5 times factor B will, with reasonable consistency of judgments, result in priorities where factor A is indeed about five times factor B.

Priorities for the alternatives with respect to scheduled maintenance are derived by calculating the principal eigenvectors of the reciprocal matrix associated with these judgments as discussed earlier. These priorities are shown in Figure 22. Because more comparisons are made (and entered into the matrix of comparisons) than are actually required to calculate the priorities, the comparison matrix is said to contain some redundancy. This redundancy is, in a sense, used to "average" errors of judgment in a manner analogous to averaging errors when estimating a population mean. The errors of judgment include errors in translating from imprecise words to the numbers that are used to represent these words in the algorithm.

OPTIMAL PATROL CRAFT

Node: 11000

Compare the relative PREFERENCE with respect to: SCHEDULE < RMA < GOAL

1=EQUAL 3=MODERATE 5=STRONG 7=VERY STRONG 9=EXTREME

1	HYDRO	9	8	7	6	5	4	3	2	**1**	2	3	4	(5)	6	7	8	9	SWATH
2	HYDRO	9	8	7	6	5	4	3	2	**1**	2	3	4	5	6	7	(8)	9	WPB 95FT
3	HYDRO	9	8	7	6	5	4	3	2	**1**	2	3	(4)	5	6	7	8	9	SES
4	HYDRO	9	8	7	6	5	4	3	2	**1**	2	3	4	(5)	6	7	8	9	PLANING
5	HYDRO	9	8	7	6	5	4	3	2	**1**	2	3	4	5	(6)	7	8	9	MONOHULL
6	SWATH	9	8	7	6	5	4	3	2	**1**	2	3	4	5	(6)	7	8	9	WPB 95FT
7	SWATH	9	8	7	6	5	(4)	3	2	**1**	2	3	4	5	6	7	8	9	SES
8	SWATH	9	8	7	6	5	4	3	2	**1**	2	3	4	5	(6)	7	8	9	PLANING
9	SWATH	9	8	7	6	5	4	3	2	**1**	2	3	4	(5)	6	7	8	9	MONOHULL
10	WPB 95FT	9	8	(7)	6	5	4	3	2	**1**	2	3	4	5	6	7	8	9	SES
11	WPB 95FT	9	8	7	6	5	(4)	3	2	**1**	2	3	4	5	6	7	8	9	PLANING
12	WPB 95FT	9	8	7	6	5	4	3	(2)	**1**	2	3	4	5	6	7	8	9	MONOHULL
13	SES	9	8	7	6	5	4	3	2	**1**	2	3	(4)	5	6	7	8	9	PLANING
14	SES	9	8	7	6	5	4	3	2	**1**	2	3	4	(5)	6	7	8	9	MONOHULL
15	PLANING	9	8	7	6	5	4	3	2	**1**	2	(3)	4	5	6	7	8	9	MONOHULL

Abbreviation	Definition
Goal	OPTIMAL PATROL CRAFT
RMA	RELIABILITY, MAINTAINABILITY, AVAILABILITY
SCHEDULE	SCHEDULED MAINTENANCE
HYDRO	HYDROFOIL
SWATH	SMALL WATER AREA TWIN HULL
WPB 95FT	WPB 95FT
SES	SURFACE EFFECT SHIP
PLANING	PLANING HULL VESSEL
MONOHULL	MONOHULL

Figure 21 – Matrix of Pairwise Comparisons

The relative preferences for the alternatives with respect to the unscheduled maintenance sub-sub-objective and the availability sub-sub-objective are

Shown in Figure 23 and Figure 24. Although the ranking of the alternatives is the same with respect to each of the three RMA sub-objectives (the 95-foot weapon patrol boat being most preferred followed by the monohull and then the planing vessel), the priorities are slightly different.

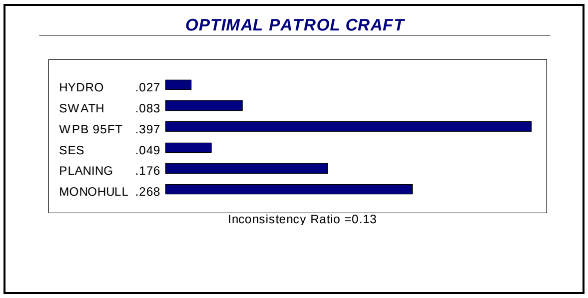


Figure 22 – Priorities with respect to Scheduled Maintenance

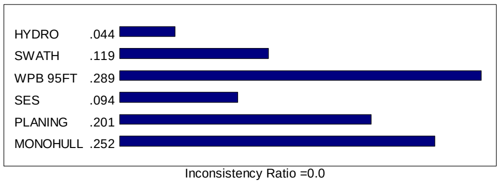


Figure 23 – Priorities with respect to Unscheduled Maintenance

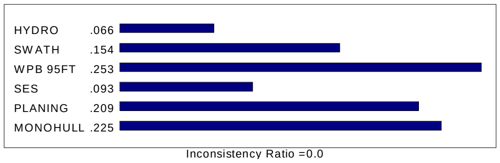


Figure 24 – Priorities with respect to availability

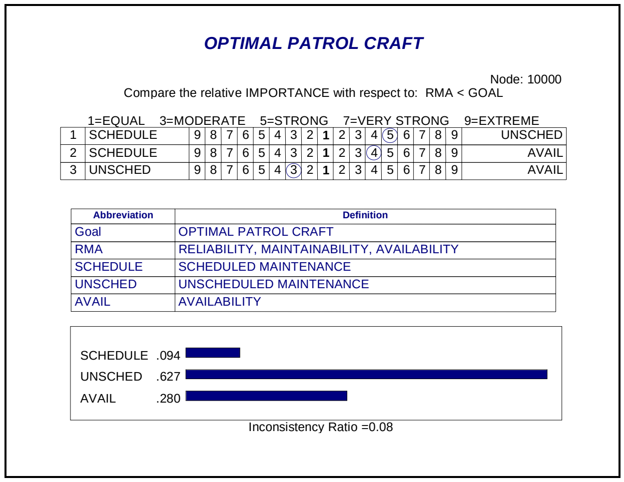


Figure 25 – Judgments and Priorities of RMA Sub-objectives

Next, the relative IMPORTANCE of the three RMA sub-objectives were determined using pairwise verbal comparisons. The judgments and resulting priorities are shown in Figure 25.

Inconsistency Ratio

The AHP comparison mode allows for inconsistent transitivity relationships and the inconsistency ratio (.08 for this last set of judgments) is an AHP measure of the lack of consistency in a set of judgments. One of the strengths of AHP is that is *does* allow for inconsistent and intransitive relationships, while, at the same time, providing a measure of the inconsistency. This strength of AHP has been criticized by multi-attribute utility theory (MAUT) and expected utility theory researchers because it does not conform to their axioms, one of which is the transitivity of preferences. Expected utility theory is grounded on the axiom of transitivity, that is, if A is preferred to B, and B is preferred to C, then A is preferred to C (or if A is three times more preferable than B and B is twice as preferable than C, then A is six times more preferable as C). For numerous reasons (discussed on page 46), the requirement that judgments be transitive and completely consistent is unrealistic. Fishburn[33] (1991, p. 130) concludes a discussion on non-transitive preferences with the following:

> Transitivity has been the cornerstone of traditional notions about order and rationality in decision theory. Three lines of research during the past few decades have tended to challenge its status. First, a variety of experiments and examples that are most often based on binary comparisons between multiple-factor alternatives suggest that reasonable people sometimes violate transitivity, and may have good reasons for doing this. Second, theoretical results show that transitivity is not essential to the existence of maximally preferred alternatives in many situations. Third, fairly elegant new models that do not presume transitivity have been developed, and sometimes axiomated, as alternatives to the less flexible traditional methods.

Luce and Raiffa[34] (1957, p. 25) discuss the transitivity axiom of expected utility theory as follows:

> No matter how intransitivities exist, we must recognize that they exist, and we can take only little comfort in the thought that they are an anathema to most of what constitutes theory in the behavioral sciences today.

They also observe:

> We may say that we are only concerned with behavior which is transitive, adding hopefully that we believe this need not always be a vacuous study. Or we may contend that the transitive description is often a 'close' approximation to reality. Or we may limit our interest to 'normative' or 'idealized' behavior in the hope that such studies will have a metatheoretic impact on more realistic studies. In order to get on, we shall be flexible and accept all of these as possible defenses, and to them add the traditional mathematician's hedge: transitive relations are far more mathematically tractable than intransitive ones.

Although one might first think that being consistent is of utmost importance, allowing for some inconsistency is reasonable. Saaty reasons that:

---

[33] Fishburn, P. C. 1991. Nontransitive Preferences in Decision Theory, Journal of Risk and Uncertainty, 4, 113-134.

[34] Luce, R.D. and H. Raiffa, 1957, *Games and Decisions*. John Wiley and Sons, Inc., New York.

The conscious mind absorbs new ideas by contrasting them through scanning or through concentration and analysis to understand how they are similar to familiar ideas. Ideas are also related to current or future activities and applied to concrete situations to test their compatibility with what is already believed to be workable. The ideas may be accepted as a consistent part of the existing understanding or they may be inconsistent with what is already known or accepted. In that case the system of understanding and practice is extended and adjusted to include the new ideas. Growth implies such expansion. If the adjustment of old ideas to accommodate a new one is drastic, then the inconsistency caused by the new idea is great. This may require considerable adjustments in the old ideas and beliefs whose old relations may no longer be intuitively recognizable and may call for a new theory or interpretation if at all possible. But such major changes cannot be made every hour, every day or even every week because it takes time to interpret and assimilate relations. Thus inconsistency arising from exposure to new ideas or situations can be threatening, unsettling and painful.

Our biology has recognized this and has developed ways to filter information in such a way that we usually make minor adjustments in what we already know when we are exposed to a new or better idea—absorbing new ideas by interpreting them from the vantage point of already established relations. Thus our emphasis on consistency exceeds our desire for exposure and readjustment. As a result, maintaining consistency is considered to be a higher priority of importance than changing. Yet the latter is also considered to be an important concern. One conclusion is that our preoccupation with consistency differs by one order of magnitude from our preoccupation with inconsistency - a split of 90% and 10%.

As discussed on page 65, perfectly consistent judgments would result in an inconsistency ratio of 0 while random judgments would, on the average, result in an inconsistency ratio of 1.0. It is possible to have an inconsistency ratio greater than 1.0, if for example, one or two judgments are accidentally inverted in what would otherwise be a very consistent set of judgments. As discussed on page 65, an inconsistency ratio of more than about 10% should alert us to the need to (1) look for and correct any clerical errors that may have caused the high inconsistency, (2) gather more information, (3) look for erroneous judgments that my have resulted from a lack of concentration, or (4) conclude that this particular aspect of the problem contains more than an average amount of inconsistency. The 10% is not an absolute rule, but rather a guideline. The more non-homogenous the factors in a cluster, the higher the inconsistency ratio is likely to be. If verbal judgments are applied to a cluster of elements that differ by more than an order of magnitude, a higher than normal inconsistency ratio can be expected (and tolerated) since the verbal scale extends only to an order of magnitude. The Numerical Matrix mode can be expanded to two orders of magnitude and the Graphical mode can accommodate two orders of magnitude.

Reducing Inconsistency

It is important that a low inconsistency not become the goal of the decision-making process. A low inconsistency is necessary but not sufficient for a good decision. It is possible to be perfectly consistent but consistently wrong. It is more important to be accurate than consistent. However, if the inconsistency ratio is higher than expected, the Expert Choice numerical matrix assessment mode menu commands can be used to locate the most inconsistent judgment, (as well as the 2^{nd} or 3^{rd}, or 9^{th} most inconsistent judgment). After locating the most inconsistent judgment in the matrix[35], the 'best fit' command will display the direction and intensity of the judgment that would be most consistent with the other judgments in the matrix. One should *not* change their judgment to this value however – but use this value to consider what might be wrong, if anything, with the judgment as it was entered – changing the judgment only if appropriate. If the judgment appears reasonable, then one should examine the 2^{nd} most inconsistent judgment and so on.

A less automated but perhaps more expedient way to examine possible inconsistencies in judgments is to use the Reorder command after calculating the priorities and then display the judgments in the numerical matrix mode[36]. Reorder changes the order of the elements in the comparison matrix according to priority (from highest to lowest). If this ranking is in fact the 'actual' ranking (assuming any erroneous judgments haven't

[35] This is only applicable to matrices consisting of four or more elements since when comparing only two factors, there can be no inconsistency. When comparing three factors, any judgment can be the most inconsistent with respect to the two remaining two judgments.

[36] If judgments were entered in the verbal or graphical mode, remember to switch back to that mode so that judgments can latter be explained and justified in the mode in which they were made.

affected the order of the elements) then there should be no inverted judgments and the judgments should be non decreasing going from left to right in each row, as well as going from bottom to top in each column.

Performance

The missions (sub-objectives), which these vessels would be required to perform, are defined by the Coast Guard Operational Program Plan and represented in the model as:

1) Search and Rescue (SAR),
2) Fisheries Conservation,
3) Drug Interdiction,
4) Marine Environmental Safety, and
5) Military Operations.

SAR – Search and rescue is an historic and extremely important mission for the Coast Guard. The service proudly proclaims itself as “The Lifesavers” and has a long tradition of protecting life and property at sea. Thus, any Coast Guard vessel must be capable of responding quickly to the reported position of a vessel in distress, often under hazardous sea conditions. The SAR capabilities of the vessel candidates in the model were judged based on a Monte Carlo simulation developed for the SAR performance evaluation.

FISHERY – Fisheries conservation is an important mission for WPBs in some operational areas. The basic function of the program is to conduct boardings of U.S. and foreign fishing fleets to ensure that their catch complies with the Coastal Fisheries Management Act. There is little requirement for speed in performing this mission but more emphasis on habitability and seakeeping. There are major fishing areas on all coasts. However, the model was developed and the comparisons made from the viewpoint of the needs of the Seventh Coast Guard District (Miami, FL) where the fisheries mission is not very important but drug interdiction is. Thus, Fisheries received a very low priority.

MARINE ENVIRONMENTAL SAFETY – MES is a mission that the Coast Guard assumed during the 1960’s with the rise of national concerns for the environment and quality of life. In general, the mission requires that virtually any available platform be prepared to respond to emergencies such as major oil or chemical spills and assist with damage control and cleanup operations. The mission is usually performed in sheltered waters near the coast where oil barges and tankers are most likely to run aground. There is some need for speed in reacting promptly to prevent the unchecked spread of an oil or chemical spill.

DRUG LAW ENFORCEMENT – Drug interdiction consumes almost 60% of the Coast Guard mission hours in the Caribbean and the Seventh Coast Guard District. There has been an enormous growth in this mission area since the escalation of the war on drugs. The performance of the candidate hull forms in the drug enforcement mission was estimated with the assistance of computer based simulations. High speed is a primary consideration in this mission area as drug smugglers turn to high performance craft to avoid coastal patrols by the U.S. Coast Guard, Navy, Customs and Drug Enforcement Agency.

MILITARY OPERATIONS – The U.S. Coast Guard, in addition to its traditional peace time duties, also constitutes a military force. In the event of a national emergency the service will be responsible for protecting the coastal waters and shores of the United States. In time of war the service becomes an operational group under the U.S. Navy and provides port security for American coastal establishments. While not a primary mission area in ordinary times, it must be considered as part of the Coast Guard's area of mission responsibilities. Priorities of the alternatives with respect to each of the performance sub-objectives and priorities for the performance sub-objectives with respect to the goal were determined using pairwise comparisons in a manner similar to that described above.

COST

Cost sub-objectives, vessel acquisition and life cycle cost were estimated by the Naval Sea System's Advanced Surface Ship Evaluation Tool (ASSET). ASSET uses a database of historical cost information to establish relationships upon which estimates for the candidate hulls are based.

LIFE CYCLE COST – Life cycle cost is the discounted cash flow of the total system over its projected life. This figure includes acquisition, personnel, maintenance, fuel and related costs.

ACQUISITION COST – Acquisition cost includes the construction cost, general and administrative (G&A), profit, spare parts inventory cost and delivery charges.

HUMAN FACTORS

HABITABILITY – The Coast Guard has long recognized that alert, comfortable personnel are far more effective in the performance of their mission duties. In some mission areas the crew is required to remain at sea for extended periods. The habitability of the vessel affects the retention and morale of skilled Coast Guard officers and crews. The habitability of the different craft can be determined subjectively by comparing the arrangement of berthing, messing, and recreational facilities.

MOTION – The ride quality of the vessel and quality of life on board are extremely important. The performance of the crew is affected by the ride of the vessel since moderate ship motion can induce seasickness and violent motion can cause injury to persons on board. Motions are defined for our model as the vessel's movements in response to the wave motion in a seaway. Motions can be empirically tested by either model or full scale testing. Motions can also be predicted by computer programs. The motions found to be most debilitating to human performance are accelerations in the vertical plane (heave) caused by the longitudinal pitching of the vessel. The judgments entered into the model were those of an expert.

BASING

Some of the advance vessels are not capable of utilizing existing Coast Guard port facilities because of draft or because they require special support services. Basing consists of hydrographic limitations and logistical concerns.

HYDROGRAPHIC LIMITATIONS – The draft and beam of candidate craft may exclude them from some locations. Dredging might not be a practical solution in some cases.

LOGISTICAL CONCERNS – If a vessel uses aluminum as a hull material it should have access in its home port area to a high strength aluminum repair facility. SWATH vessels must be located near a facility capable of dry docking such an unusual hull form.

Judgments for alternative preference with respect to sub-objectives below each of the other major objectives and for the relative importance of the sub-objectives were performed in a similar manner as that descried for the RMA objective above. Finally, verbal judgments were made about the relative importance of the major objectives. These judgments and the resulting priorities are shown in Figure 26 & Figure 27.

OPTIMAL PATROL CRAFT

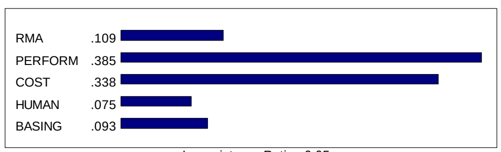
Node: 0

Compare the relative IMPORTANCE with respect to: GOAL

1=EQUAL 3=MODERATE 5=STRONG 7=VERY STRONG 9=EXTREME

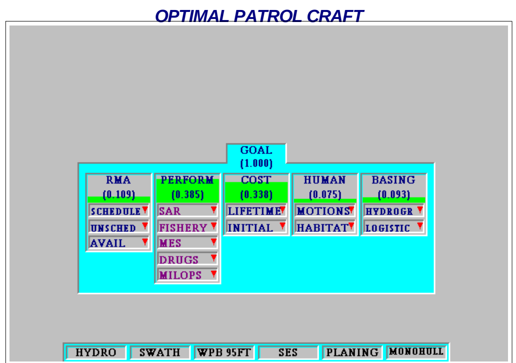
1	RMA	9 8 7 6 5 4 3 2 **1** 2 3 4 (5) 6 7 8 9	PERFORM
2	RMA	9 8 7 6 5 4 3 2 **1** 2 3 4 (5) 6 7 8 9	COST
3	RMA	9 8 7 6 5 4 (3) 2 **1** 2 3 4 5 6 7 8 9	HUMAN
4	RMA	9 8 7 6 5 4 3 2 (**1**) 2 3 4 5 6 7 8 9	BASING
5	PERFORM	9 8 7 6 5 4 3 2 (**1**) 2 3 4 5 6 7 8 9	COST
6	PERFORM	9 8 7 6 (5) 4 3 2 **1** 2 3 4 5 6 7 8 9	HUMAN
7	PERFORM	9 8 7 6 5 (4) 3 2 **1** 2 3 4 5 6 7 8 9	BASING
8	COST	9 8 7 6 5 4 (3) 2 **1** 2 3 4 5 6 7 8 9	HUMAN
9	COST	9 8 7 6 5 4 (3) 2 **1** 2 3 4 5 6 7 8 9	BASING
10	HUMAN	9 8 7 6 5 4 3 2 (**1**) 2 3 4 5 6 7 8 9	BASING

Abbreviation	Definition
Goal	OPTIMAL PATROL CRAFT
RMA	RELIABILITY, MAINTAINABILITY, AVAILABILITY
PERFORM	MISSION PERFORMANCE
COST	COST
HUMAN	HUMAN FACTORS
BASING	HOME PORT CRITERIA



Inconsistency Ratio =0.05

Figure 26 – Judgments and Priorities for Major Objectives



Abbreviation	Definition
AVAIL	AVAILABILITY
BASING	HOME PORT CRITERIA
COST	COST
DRUGS	DRUG INTERDICTION
FISHERY	FISHERY ENFORCEMENT
HABITAT	HABITABILITY
HUMAN	HUMAN FACTORS
HYDROGR	HYDROGRAPHICS - WATER DEPTH
INITIAL	ACQUISITION COSTS
LIFETIME	LIFE CYCLE COSTS
LOGISTIC	LOGISTIC
MES	MARINE ENVIRONMENT SAFETY
MILOPS	MILITARY OPERATIONS
MOTIONS	SEA STATE MOTIONS
PERFORM	MISSION PERFORMANCE
RMA	RELIABILITY, MAINTAINABILITY, AVAILABILITY
SAR	SEARCH AND RESCUE
SCHEDULE	SCHEDULED MAINTENANCE
UNSCHED	UNSCHEDULED MAINTENANCE

(Priorities shown are 'Local'

Figure 27 – Priorities from Goal node

Having decomposed the complex problem into its constituent parts, and having derived priorities, a synthesis was performed. The synthesis details are shown in Figure 28, and Figure 29. The results are shown in Figure 30.

OPTIMAL PATROL CRAFT

Synthesis of Leaf Nodes with respect to GOAL
Ideal Mode
OVERALL INCONSISTENCY INDEX = 0.03

LEVEL 1	LEVEL 2	LEVEL 3	LEVEL 4	LEVEL 5
RMA =.109				
	SCHEDULE=.010			
		HYDRO <.001		
		SWATH =.002		
		WPB 95FT=.010		
		SES =.001		
		PLANING =.005		
		MONOHULL=.007		
	UNSCHED =.068			
		HYDRO =.010		
		SWATH =.028		
		WPB 95FT=.068		
		SES =.022		
		PLANING =.048		
		MONOHULL=.060		
	AVAIL =.031			
		HYDRO =.008		
		SWATH =.019		
		WPB 95FT=.031		
		SES =.011		
		PLANING =.025		
		MONOHULL=.027		
PERFORM =.385				
	SAR =.114			
		HYDRO =.080		
		SWATH =.019		
		WPB 95FT=.025		
		SES =.018		
		PLANING =.114		
		MONOHULL=.019		
	FISHERY =.000			
		HYDRO =.000		
		SWATH =.000		
		WPB 95FT=.000		
		SES =.000		
		PLANING =.000		
		MONOHULL=.000		
	MES =.023			
		HYDRO =.023		

Figure 28 – Synthesis details

OPTIMAL PATROL CRAFT

		SWATH =.023		
		WPB 95FT=.023		
		SES =.023		
		PLANING =.023		
		MONOHULL=.023		
	DRUGS =.170			
		HYDRO =.096		
		SWATH =.036		
		WPB 95FT=.046		
		SES =.089		
		PLANING =.170		
		MONOHULL=.101		
	MILOPS =.077			
		HYDRO =.077		
		SWATH =.070		
		WPB 95FT=.015		
		SES =.042		
		PLANING =.055		
		MONOHULL=.048		
COST =.338				
	LIFETIME=.281			
		HYDRO =.114		
		SWATH =.157		
		WPB 95FT=.281		
		SES =.161		
		PLANING =.167		
		MONOHULL=.185		
	INITIAL =.058			
		HYDRO =.011		
		SWATH =.026		
		WPB 95FT=.058		
		SES =.028		
		PLANING =.031		
		MONOHULL=.037		
HUMAN =.075				
	MOTIONS =.044			
		HYDRO =.043		
		SWATH =.044		
		WPB 95FT=.004		
		SES =.014		
		PLANING =.009		
		MONOHULL=.006		
	HABITAT =.031			
		HYDRO =.031		

Figure 29 – Synthesis details continued

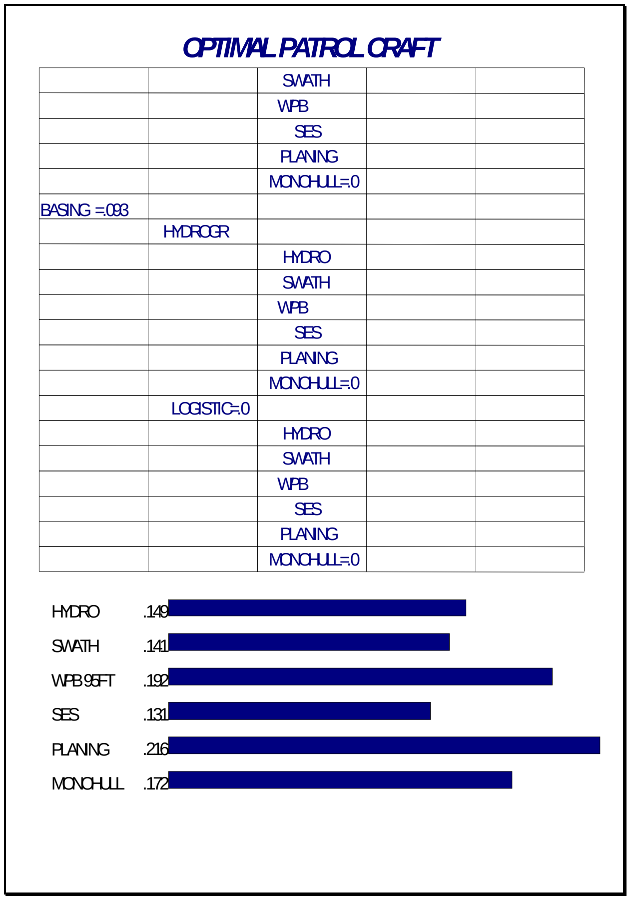


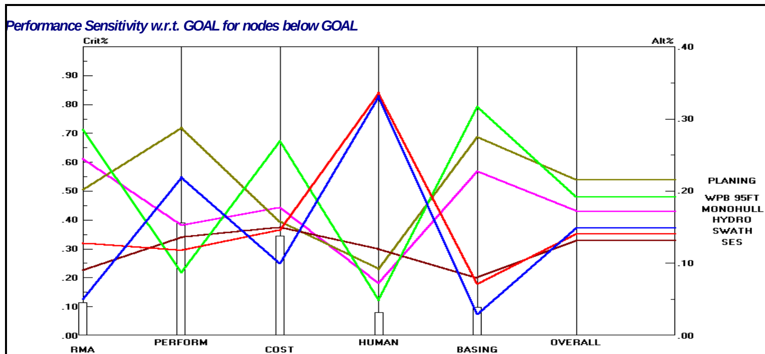
Figure 30 – Synthesis Results

RESULT

The planing hull is the best choice, based on the subjective judgments, empirical data, and simulated outputs entered for the designated operating area. The results of the model were also intuitively appealing[37]

Sensitivity Analysis

A performance sensitivity graph (Figure 31) shows how well each alternative performs with respect to each of the major objectives. The importance of the objectives are depicted by vertical bars[38].



Abbreviation	Definition
RMA	RELIABILITY, MAINTAINABILITY, AVAILABILITY
PERFORM	MISSION PERFORMANCE
COST	COST
HUMAN	HUMAN FACTORS
BASING	HOME PORT CRITERIA

HYDRO	HYDROFOIL
SWATH	SMALL WATER AREA TWIN HULL
WPB 95FT	WPB 95FT
SES	SURFACE EFFECT SHIP
PLANING	PLANING HULL VESSEL
MONOHULL	MONOHULL

Figure 31 – Performance Sensitivity

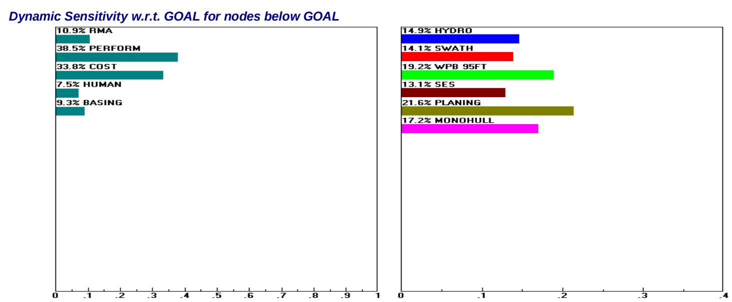
[37] If the results were not intuitive, then asking *why not* would lead to a modification of the model, and/or judgments and/or a change in intuition.

[38] Priorities for the objectives can be read from the left-hand axis.

The performance of each alternative with respect to each of the objectives is depicted by the intersection[39] of a colored sequence of line segments with the vertical line at each of the objectives. The overall performance of an alternative is depicted by the intersection of the alternative's line segments with the "overall" vertical line at the right of the graph[40].

Ordinal, interval *and* ratio information is conveyed in all of the priorities. For example, the planing hull vessel, depicted by the blue sequence of line segments, is third best for RMA, best for performance, third best for cost, fourth best for human factors, second best for basing, and best overall. But in addition to knowing the order of performance for each alternative with respect to each objective as well as overall, we also know and can see the intervals and ratios between alternatives. Looking at the human factors objective, for example, the interval between the first and second alternatives (SWATH and HYDRO) is insignificant, whereas the interval between the second and third alternatives (HYDRO and SES) is very large. The interval between the best and second best overall alternatives (planing hull and WBP) is much larger with respect to performance than with respect to basing. This is useful information. (Measurement to at least an interval level is necessary in order to determine which alternative is best overall. It would be incorrect to conclude that the WPB was best overall because it was best on three of the five objectives).

[39] Performance of the alternatives can be read from the right hand axis.
[40] Alternative line segments between the objectives are drawn for visual clarity only and do not convey any information.



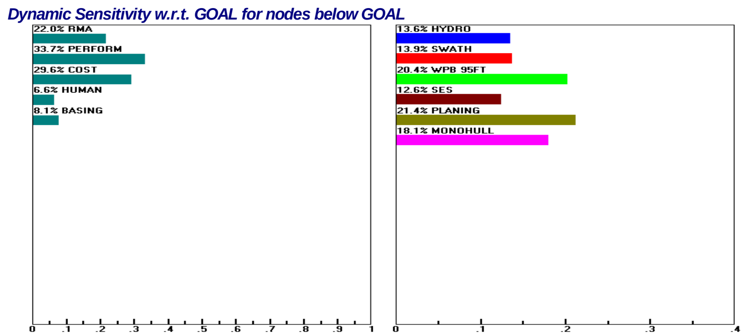
Abbreviation	Definition
RMA	RELIABILITY, MAINTAINABILITY, AVAILABILITY
PERFORM	MISSION PERFORMANCE
COST	COST
HUMAN	HUMAN FACTORS
BASING	HOME PORT CRITERIA

HYDRO	HYDROFOIL
SWATH	SMALL WATER AREA TWIN HULL
WPB 95FT	WPB 95FT
SES	SURFACE EFFECT SHIP
PLANING	PLANING HULL VESSEL
MONOHULL	MONOHULL

Figure 32 – Dynamic Sensitivity

Ratio interpretations are also meaningful. For example, with respect to RMA, WPB is about six times more preferable than the hydrofoil. Or, for example, the overall priority of the most preferred alternative, the planing hull vessel (about 22%), is not even twice that of the least preferred alternative, the SES (about 13%). A dynamic sensitivity graph is shown in Figure 32.

Suppose we thought that the importance of RMA might be, or become, more important than the 10.9% shown. By 'pulling' the RMA bar to the right, each of the other bars decrease (in proportion to their original priorities) and the overall priorities of the alternatives change as well. The dynamic sensitivity graph with RMA increased to about double its original value is shown in Figure 33. The overall priority of the planning vessel has decreased while priority of the weapon patron boat has increased – however the planing hull vessel is still slightly preferred. Thus, the current result is relatively insensitive to even a doubling in priority of the RMA objective.



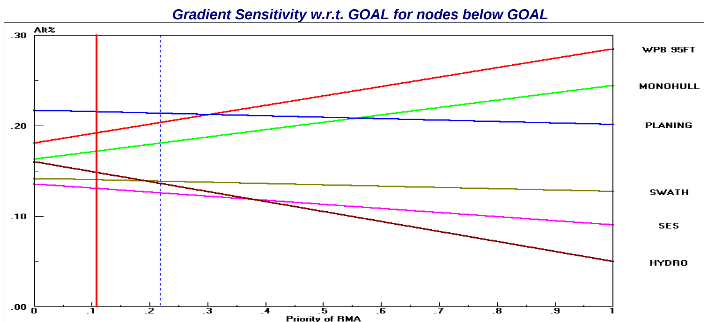
Abbreviation	Definition
RMA	RELIABILITY, MAINTAINABILITY, AVAILABILITY
PERFORM	MISSION PERFORMANCE
COST	COST
HUMAN	HUMAN FACTORS
BASING	HOME PORT CRITERIA

HYDRO	HYDROFOIL
SWATH	SMALL WATER AREA TWIN HULL
WPB 95FT	WPB 95FT
SES	SURFACE EFFECT SHIP
PLANING	PLANING HULL VESSEL
MONOHULL	MONOHULL

Ideal Mode

Figure 33 – After doubling RMA

A gradient sensitivity graph for the RMA objective is shown in Figure 34. The red vertical line shows that the original priority of RMA was about 11%. The gradient (slope) of each of the alternative lines (blue for RMA, red for WPB, etc.,) represents the rate of change in priority for the alternative as the priority of RMA is changed. Thus, for example, the priority of the planing hull vessel would decrease while that of the weapon patrol boat would increase if the priority of the RMA objective is increased (red bar is moved to the right.) A change in one sensitivity plot causes changes to each of the other sensitivity plots. The blue dotted vertical line in the gradient sensitivity plot (Figure 34) corresponds to the doubling of the importance of the RMA objective on the dynamic sensitivity graph (Figure 33). The planing hull vessel is still preferred to the weapon patrol boat at this RMA priority. However, if RMA were increased to beyond about 30.3%, the weapon patrol boat would become the preferred alternative. If the priority of RMA were increased to 100% (the vertical bar moved all the way to the right, the priorities would correspond to the intersection of each alternative's line segment with the vertical line for the RMA objective on the performance sensitivity graph (Figure 31). That is, if RMA were the only objective then the planing hull vessel would drop to the third most preferred alternative.



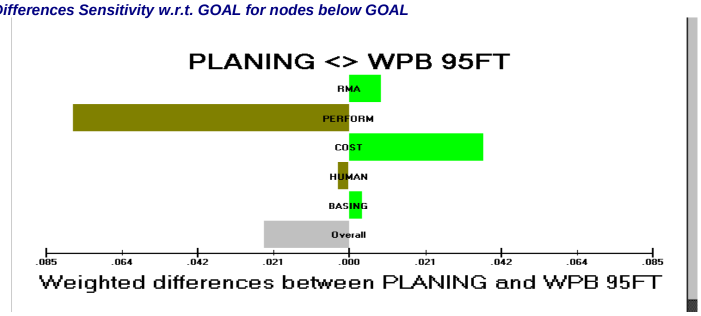
Abbreviation	Definition
RMA	RELIABILITY, MAINTAINABILITY, AVAILABILITY
PERFORM	MISSION PERFORMANCE
COST	COST
HUMAN	HUMAN FACTORS
BASING	HOME PORT CRITERIA

PLANING	PLANING HULL VESSEL
WPB 95FT	WPB 95FT
MONOHULL	MONOHULL
HYDRO	HYDROFOIL
SWATH	SMALL WATER AREA TWIN HULL
SES	SURFACE EFFECT SHIP

Figure 34 – Gradient Sensitivity

A graph of the differences between the top two alternatives is shown in Figure 35. The planing hull vessel is preferred to the weapon patrol boat on performance by quite a margin, and only slightly on human factors, while the weapon patrol boat is preferred to the planing null vessel on cost, less so on RMA and even less on basing.

Differences Sensitivity w.r.t. GOAL for nodes below GOAL



Abbreviation	Definition
RMA	RELIABILITY, MAINTAINABILITY, AVAILABILITY
PERFORM	MISSION PERFORMANCE
COST	COST
HUMAN	HUMAN FACTORS
BASING	HOME PORT CRITERIA

HYDRO	HYDROFOIL
SWATH	SMALL WATER AREA TWIN HULL
WPB 95FT	WPB 95FT
SES	SURFACE EFFECT SHIP
PLANING	PLANING HULL VESSEL
MONOHULL	MONOHULL

Ideal Mode

Figure 35 – Difference Sensitivity

Seven Step Process for Choice

We have now looked at four steps in applying AHP and Expert Choice to a choice problem: (1) decompose the problem – develop a hierarchy; (2) establish priorities; (3) synthesis; and (4) sensitivity analysis. These steps can be elaborated by imbedding them in a more encompassing seven step process as follows:

Step 1: Problem definition and research

1a: Problem *identification*

1b: Identify *objectives*[41] and *alternatives*. A listing of the *pros and cons* of each alternative is often helpful in identifying the objectives.

---

[41]A rational decision is one made on the basis of *objectives*. *Criteria* are used to measure how well we achieve our objectives. The words objectives and criteria, although having different definitions, are commonly used interchangeably in analyzing a decision. The word *objective* is preferred because it helps us

1c: Research the alternatives

Step 2: Eliminate infeasible alternatives

2a: Determine the *"musts"*

2b: Eliminate alternatives that do not satisfy the "musts"

Step 3: Structure a decision model in the form of a hierarchy to include goal, objectives (and sub objectives), and alternatives. Add other factors (such as actors and scenarios) as required

Step 4: Evaluate the factors in the model by making pairwise relative comparisons

4a: Use as much factual data as is available, but remember to *interpret* the data as it relates to satisfying your objectives. (That is, do not assume a linear *utility curve* without thinking about whether it is a reasonable assumption).

4b: Use knowledge, experience, and intuition for those qualitative aspects of the problem or when hard data is available.

Step 5: Synthesize to identify the "best" alternative

Step 6: Examine and verify decision, iterate as required

6a: Examine the solution and perform sensitivity analyses. If the solution is sensitive to factors in the model for which you do not have the best data available, consider spending the time and money to collect the necessary data and iterate back to step 4.

6b: Check the decision against your *intuition*. If they don't agree, ask yourself why your intuition tells you a different alternative is best. See if the reason(s) are already in the model. If not, revise the model (and or judgments). Iterate as required. In general you will find that both your model

---

focus on what we are trying to achieve as we perform the evaluation. However, the word criterion is commonly used in practice.

and intuition may change (i.e. you are learning). When your intuition (possibly different now than it was before) and the model agree, continue on to step 7.

Step 7: Document the decision for justification and control

Other decision-making 'recipes'

Numerous other recipes for decision-making have been published over the years. Most have similar components. One of the most widely disseminated recipes is the following seven steps from Kepner &Trego[42]

1) establish objectives
2) classify importance of objectives
3) develop alternative actions
4) evaluate alternatives against established objectives
   determine musts
   determine wants
5) choose the alternative that is best able to achieve the objectives as tentative decision
6) explore the tentative decision for future possible adverse consequences
7) control the effects of the final decision by taking other actions to prevent possible adverse consequences from becoming problems, and by making sure the actions are carried out.

Musts and Wants

Each of the above 'recipes includes the considerations of 'musts'. Kepner &Trego advised using 'musts and wants' in their step 4. 'Musts' are used to eliminate infeasible alternatives in step 2 of our 'recipe.' An alternative that does not satisfy any one (or more) of the musts is said to be 'infeasible' and is eliminated from consideration. The remaining

---

[42] C. H. Kepner and B. B. Tregoe, B.B. *The Rational Manager: A Systematic Approach to Problem Solving and Decision Making*, McGraw Hill, New York, NY, 1965.

alternatives are then evaluated based on how well they meet the 'wants'. We can look at 'musts' in terms of Simon's bounded rationality and Janis' common simplistic strategies discussed earlier. Musts represent aspiration levels that are used in a satisficing mode of decision-making. However, instead of selecting the first alternative that meets all of our 'musts' as is done as a common simplistic strategy when satisficing, we continue to identify additional alternatives that also satisfy our musts and then make a selection based on how well the 'feasible' alternatives satisfy our 'wants'. The 'wants' are the *objectives* that form the basis of our definition of what is a rational choice.

Attributes, such as the cost of an automobile, or the price to earnings ratio of a stock, can be *both* musts and a wants. For example, in selecting an automobile, we might require that the automobile must cost less than $40,000 (eliminating all those that cost more than that) and then use (low) cost as a want (one of our objectives) in evaluating the remaining, feasible, alternatives. We make the following recommendations when considering 'musts':

Do not use 'musts' alone as a way of making any important decision. For example, don't offer to buy the first house that you find that meets your 'musts'.

It is often to use 'musts' to narrow down the alternatives before making a compensatory evaluation. The advantage of doing this is that it is relatively easy to do and it reduces the complexity of the compensatory evaluation. The disadvantage is that it is possible to be shortsighted and eliminate one or more alternatives that, in the compensatory evaluation, might have been preferred because of the degree to which one or more of its other attributes contribute to other objectives.

If there are too many alternatives to evaluate in a practical sense, (such as stocks on the major exchanges), uses musts (such as stocks must have a price to earnings ratio of not more than 40 to 1 and a capitalization of at lease $1 Billion) to reduce the alternatives to a reasonable number. However, try to be conservative in specifying the aspiration levels for the musts.

Summary of the benefits of AHP

In summary, AHP makes it possible for executives to assimilate all the facts, weigh the pluses and minuses, reach, re-evaluate, and communicate their decision. Once an initial decision is made, it is not final; even a strong willed decision maker is subject to external pressures from special interest groups such as, suppliers, customers, employees, trade unions or politicians. Objectives that were thought to be central to a decision, may, under these outside influences, become less central or dominant and a re-evaluation becomes necessary. Gradually, priorities are changed until a new, re-shaped, decision emerges. Without a decision model audit-trails are lost and executives find it impossible to systematically review or retrace the steps and sub-decisions made in the decision process. The difficulty of conducting a proper review increases exponentially with the number of objectives.

Incremental Improvement

Now that we have seen some of how AHP works, we address the question of how you might begin using AHP in your organization. There are two basic approaches corresponding to two important business processes that have been widely publicized: (1) business process re-engineering, and (2) incremental improvement. Re-engineering is typically more costly and risky and requires commitment and freedom from constraints that are more the exception than the rule. Continuous improvement is often an easier starting path. By examining how your organization currently makes decisions or performs evaluations, and asking how the details of the current process(es) fair relative to the decision-making concepts embodied in AHP, you will see many opportunities for easy, yet significant improvements. After a few such improvements, people in your organization should become ‘comfortable’ with the changes and perhaps willing to make a commitment to a re-engineering of the organizations basic decision processes.

Retirement Places Rated

We will illustrate how an existing 'successful' evaluation methodology, rating retirement places, can be examined, found 'lacking' in one or more aspects, and 'improved' using some of the concepts on which AHP is based.

Background

There are many kinds of location decisions. One such decision that is becoming more widespread is selecting a location for retirement. The best reference we found on this topic is a book by David Savageau called *Retirement Places Rated*[43]. Although the former head of the National Institute on Aging, Dr. Robert Butler, advised that the best place to retire for most people "is the neighborhood where [they] spent [their lives]", Savageau says "possibly,— just possibly—there is someplace in this country where you might prosper more than the place where you now live". Perhaps Savageau recognized that as our society has become more mobile, fewer and fewer people spend the majority of their lives in any one place, so perhaps the need to select a retirement place has become more important than in the past. To take advantage of this 'opportunity', Savageau collected interesting and useful information about more than one hundred fifty retirement places throughout the United States. Realizing that just presenting this information in a book was not enough, Savageau graded the places on the basis of seven[44] factors that he felt influences the quality of retirement life: money matters, climate, job outlook, available services, housing, leisure, and personal safety.

Savageau recognized the difficulty in rating retirement places. Early in his book he states:

> "There are three points of view on rating places. The first says that defining what's good for all people at all times is not only unfair, it's impossible and shouldn't be tried at all. Another view says you can but shouldn't because measuring a touchy thing like "livability" pits cities and towns against each other and leads to wrong conclusions. The third point of view says do it as long as you make clear what

---

[43] Savageau, David, *Retirement Places Rated,* Macmillan, New York, N.Y., 1995
[44] Notice the 'magic number seven' appearing again.

> your statistical yardsticks are and go on to use them consistently. Although the first and second positions may be valid, Retirement Places Rated sides with the third.

Even though *Retirement Places Rated* is the best book we found on the subject, and we agree that there is a need to rate, score, evaluate – call it what you will – in order to make a decision, there is considerable room for improvement in the concepts and methods contained in *Retirement Places Rated*. Just making your statistical yardsticks clear and using them consistently is not enough! In all fairness to Savageau, much of what we address below as inadequacies (opportunities for improvement) in *Retirement Places Rated* and its methodology are difficult to circumvent given the constraints of a static book. However, these constraints are easy to circumvent in today's world of personal computer availability. People making the important decision about where to retire need not settle for what we can best describe as 'inferior' decision-making. We next look at the following opportunities for incremental improvement:

- measuring and synthesizing from ratio rather than ordinal scales
- structuring hierarchically to better understand relationships and avoid double counting
- deriving meaningful priorities of objectives

Change in methodology – from adding ranks to averaging scores

Retirement Places Rated has more data than most people would ever care to peruse[45]. Savageau had to find ways to help interpret and synthesize this data because its impossible to synthesize it all mentally[46]. Savageau did this with the best tools he knew how to use, creatively "inventing" and refining formulas and methodologies as he progressed. In the 1990 edition of his book, Savageau ranked 151 retirement places on each of seven criteria and then added the ranks[47] to determine an overall ranking. Using this methodology, Las Vegas Nevada came in 105th. If we review the decision-

[45] Although our Murphy's axiom of data "the data you have you don't need and the data you need you don't have" is sure to apply in part.
[46] See page 6.
[47] Adding ranks, is as we now know, a mathematically meaningless operation since ranks are 'ordinal' level data and cannot be meaningfully added.

making concepts presented in an earlier section of this book, we will observe that the addition of ranks is mathematically meaningless! Perhaps Savageau recognized this himself, even though his book was commercially successful.

In the next (1994) edition of *Retirement Places Rated* Savageau improved his methodology by 'scoring' (instead of ranking) 183 cities on the seven criteria and then averaging the scores to determine an overall score. If the scores possess interval or ratio meaning (and presumably they do), then this is an obvious improvement. This time, Las Vegas came in first! A casual observer might conclude that Las Vegas must have improved remarkably in those five years. But the change in the result is due much more to the change in methodology than any changes in Las Vegas. Looking at Table 4 we see that Las Vegas' ranks improved in some categories, but actually got worse in others. The total of the ranks for Las Vegas did improve from 521 to 504 between the two editions of the book. But how could this small change in the rank total cause Las Vegas's overall standing to change from 105th to 1st? The answer is that adding ranks was (and is) mathematically meaningless so it is not too surprising when we see things like this happen.[48] We will now look more closely at Savageau's methodology and how it can be further improved.

Table 4 – Ranks of Las Vegas

Year Rated	1990	1995
Money Matters	128	117
Housing	128	86
Climate	63	38
Personal Safety	154	140
Services	30	105
Working	5	8
Leisure Living	13	10
Total Ranks	521	504

[48] We often wonder how often misleading results due to the adding of ranks are produced but go unrecognized!

An Expert Choice representation of Savageau's 'model', along with subcriteria he included under some of the criteria, is shown in Figure 36

Creating 'magic' formulas

To avoid adding ranks, Savageau needed a way to 'score' the cities on each of his lowest level 'criteria'[49] or subcriteria. To do this, he creatively invented ways to translate data into 'scores'. For example,. when evaluating the cities with respect to the Housing criterion, Savageau reasoned that the sum of mortgage payments and property taxes, as a percentage of local household income, was a reasonable measure to use. Lacking a scale on which to place these measurements (as well comparable scales for measurements for the other criteria and sub-criteria), he did what he could – he 'made up' formulas for constructing scales[50]. His housing scale, for example, is constructed using the following formula:

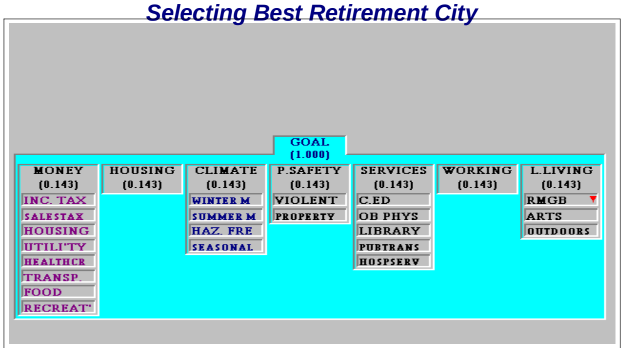
> Mortgage payments and property taxes are added together and expressed as a percent of local household income. This percent figure is then graded against a standard where half the typical 25% mortgage lending requirement gets 100 and three times the 25% lending requirement gets a 0.

---

[49] We will call these 'objectives' to make the evaluation more meaningful.

[50] A common misconception and one made in *Retirement Places Rated*, is that if each of several separate scales have the same 'range', in this case 0 to 100, then they are 'comparable' and can be added. This is not necessarily true!

Selecting Best Retirement City



Abbreviation	Definition
ARTS	Lively arts
C.ED	Continuing Education
CLIMATE	Climate
FOOD	Cost of Food
HAZ. FRE	Hazard free
HEALTHCR	Health Care Costs
HOSPSERV	Short-term Hospital Services
HOUSING	Home mortgage payments and property taxes
INC. TAX	State Income Taxes
L.LIVING	Leisure Living
LIBRARY	Public libraries
MONEY	Money matters
OB PHYS	Office based Physicians
OUTDOORS	Outdoor Assets
P.SAFETY	Personal Safety
PROPERTY	Property Crimes
PUBTRANS	Public transportation
RECREAT'	Recreation Costs
RMGB	Restaurants, Movies, Golf, Bowling
SALESTAX	State and local sales taxes
SEASONAL	Seasonal affect
SERVICES	Services
SUMMER M	Summer mildness

Figure 36 – Criteria and subcriteria

This formula, although perhaps as good as any other, is rather arbitrary! Why does a location with a percent equal to ‘half’ the typical 25% get a score of 100? Why not score it 100 if the percent is 1/10th the typical 25% instead? Even if we assume that those using the Savageau’s book agree with basing the score on a percentage of local income, by using Savageau’s formula, a location with a (mortgage payment plus property tax percent of local household income) value of 2.5% would get the same score as another location with a value of 12.5%. Obviously, they should not be equal! But is the former five times more preferable? Perhaps not, since the ‘utility’ of lower payments may not be linear for any individual or couple. The point is, there is no magic formula that will fit everyone. All we can say that in general, the scale is subjective, and probably non-linear at that.

The Climate criterion has four subcriteria, as seen in Figure 36 Savageau's formula for Winter Mildness is:

Winter Mildness: Equal weights to:

(a) Winter severity –average apparent (influenced by wind chill) temperature from November through April

scale: 0 = 0; 55=100

(b) Winter length – # of days when temperature falls below freezing

scale: 0 = 100; 365=0

While this 'formula' might, on the surface, seem reasonable, a look at some of the resulting values shows that it does not convey what most people looking for a retirement location would feel. The ratio of preferences for Florida over Maine for winter mildness (see Table 7) using Savageau's data and formula is *less* than two to one! Preferences can be estimated in many ways[51]. Whatever the way, common sense tells us that the ratio of values for the preference of Florida over Maine (for winter mildness) should be *much more* than two to one and thus Savageau's formula does *not* provide a good measure! It's possible, although not practical in the context of a hard copy book like *Retirement Places Rated*, to improve on such 'magic' formulas based on 'objective' data. Fortunately, we are no longer limited to hard copy books alone. Using the Analytic Hierarchy Process and readily available computer technology, we can elicit judgments from decision maker(s) about how many times more preferable Florida is than Maine for winter mildness. Pairwise comparisons can be made verbally, numerically or graphically. A sequence of pairwise comparisons, can, as we have seen above with Expert Choice, derive accurate ratio scale measures of

[51] Economists often take the approach, which we don't fully subscribe to -- to convert everything to 'dollars'. In this context, an economist might compute the ratio of the average winter hotel room rental rates in Bar Harber Maine vs. Lakeland Florida (adjusting, as Economists try to do) for other factors such as differences in economic conditions of the two areas. Another approach an economist might use would be to estimate the ratio of the number of retired people going 'South' for the winter rather than 'North' for the winter. But this measure too would have to somehow be adjusted for other related factors, not indicative of 'winter mildness' such as those who go to Maine in the winter in order to Ski.

preference for the alternatives for each objective (criterion), which in turn, can be meaningfully synthesized.

Savageau justifies his methodology by saying "do it as long as you make clear what your statistical yardsticks are and go on to use them consistently." Assuming things are correct because they are consistent is a common misconception. Can we conclude that since Savageau's Summer Mildness formula (see Table 5) produces substantially the same results (although in the opposite direction) as the winter formula, then together they form a consistent approach? The answer is obviously No. Two wrongs do not make a right. Consistency[52] is a necessary but not sufficient condition for correctness.

Table 5 – Savageau's Summer Mildness formula

Location	Winter Mildness	Summer Mildness
Bar Harbor ME	56	94
Lakeland-Winter Haven FL	100	53

Double Counting

Notice that 'Housing' appears in two places in Figure 36: both as a subcriterion of money matters, and as a criterion on its own. True, it does 'fit' both places, but we may be counting the same thing twice. If we focus on 'objectives', rather than 'criteria' we can more easily see whether this is or is not the case. What do we mean when we say housing is a criterion? The answer is not clear – only that we will use housing somehow in evaluating the alternatives. However, when we say that housing is an *objective*, and ask ourselves what we would *want* in terms of housing, several things come clearly to mind – cost of housing, availability of housing, age of housing, style of housing, etc. By including housing as both a subcriterion below money matters and as a criterion on its own, and by using formulas that score alternative locations based on the cost of housing

[52] A reasonable amount of consistency is necessary, but perfect consistency is not necessary.

in each instance[53], this model counts the same *objective* twice! The severity of the double counting depends on how much 'weight' is 'given' to the main housing criterion and the housing criterion below money matters. In the 'Putting it All Together' chapter of his book, Savageau uses equal weights for all the major criteria[54], with about 45% of the money matters weight going again to the housing sub-criterion! This is a considerable amount of double counting.

A better 'model' would be to have 'cost of housing' as a sub-objective under money matters as it now stands, and eliminate the stand alone housing criterion. Still better would be to include a major criterion representing the non-monetary housing factors, such as availability of housing styles[55]. Structuring the evaluation as a hierarchy and focusing on objectives rather than criteria will give us a much better idea of what we are evaluating in our 'model'. The model can help us to synthesize the multitude of factors involved in the evaluation by enabling us to determine how well the alternatives perform relative to our objectives!

Equal Weighting

As mentioned above, Savageau used equal weights for each of the seven criteria in calculating an overall score for the alternative retirement locations. However, he recognized that this wasn't necessarily the best thing to do. Savageau explained:

> "At the end of [this] book, in the chapter entitled "Putting It All Together", money matters, housing, climate, personal safety, community services, working and leisure living get equal weight to identify retirement places with across-the-board strengths.

---

[53] Savageau's formula for scoring locations based on housing costs under money matters is based on the cost of housing index for each location: housing costs that are 25 percent below the national average get a 100 and housing costs double the national average get a 0. This formula, while different from the one used for the criterion housing under the goal, not only uses information based on the same costs, but obviously points to the same objective -- lower housing costs.

[54] What else could he do? The weights are subjective and will depend on who is deciding where to retire.

[55] Savageau does indeed include information about historic neighborhoods, but these are not considered in deriving the scores for the alternative locations.

> You may not agree with this system. To identify which factors are more important and which factors are less, you might want to take stock of your own preferences."

Toward that end, Savageau provides a "Preference Inventory" consisting of sixty three pairwise comparisons of the following form:
For each numbered item, decide which of the two statements is more important to you when choosing a place to retire.

A) The cost of living, or
B) Historic homes in an area
C) The duration of the winter, or
D) The odds of being a crime victim., etc.

Savageau then says to "count all the marks you've made in the boxes next to the letter A, and write that down next to money-matters; similarly items with the letter B correspond to housing, etc."

This is a valiant attempt to move away from equal weighting, which most people agree is not appropriate. However Peter Drucker emphasized (see page 45) we must 'measure', not count. Let's see why counting is not very accurate. There are 21 possible pairs of the seven criteria (or the letters A through G). Savageau provided three questions for every possible pair (63 questions in all), with each question containing a different aspect of the respective criterion. Lets consider just 21 of the questions[56] and a question representing each criterion will appear in six of these questions. Suppose someone's 'true' weights as determined from the principle eigenvector of a matrix of pairwise comparisons are as shown in the second column of the following table[57]:

[56] The following results will be the same if we consider all 63 questions.
[57] The following arguments hold regardless of whether the true order is A, B, C, D, E, F, G or any other order.

Table 6 – True Weights

	True Weights	Count	Calculated
A	45	6	28.6
B	16	5	23.8
C	14	4	19.0
D	10	3	14.3
E	8	2	9.5
F	5	1	4.8
G	2	0	0

Thus A is preferred to B, B to C and so on. If we assume assessments of the 21 questions that are consistent with the order of priorities in column 1, A will be preferred in each of the six questions in which it appears. That is A will be preferred in the question in which it is paired with B, and in the question in which it is paired with C, and so on. A will be preferred 6 times. B will be preferred in questions in which it is paired with C, D, E, F and G, or five times. C will be preferred in questions in which it is paired with D, E, F, and G, or four times. And so on. The count of the number of times each criterion is chosen and the resulting weights are shown in columns three and four of the table. Assuming accurate and consistent judgments, a counting technique like this would produce the same calculated weights *regardless* of the 'true' weights. Thus, for this example, the weight for criterion A, money matters, is much less than the 'true' weight. Generalizing, we can say that this counting technique will produce incorrect results in *all* cases except when the 'true' weights happen to be proportional with ratios of 6:5:4:3:2:1.

Another difficulty with this technique is the use, in the pairwise comparison questions, of different dimensions of a criterion as surrogates for the criterion. A criterion with five subcriteria, one of which is much more important than the others, would receive far less weight than it deserves because it would 'win' the comparison in the question containing an aspect of the most important sub-criterion, but might 'lose' the comparisons in questions involving the less important sub-criteria.[58]

Incremental improvement using AHP

After seeing how AHP can be used to measure and synthesize, how could you incrementally improve on the methodology used in *Retirement Places Rated*? Some of the possibilities include:

- Focus on objectives rather than criteria. This will make clear when double counting is inappropriate (and, in some cases, when it is appropriate).
- Have the decision maker(s) derive priorities for the relative importance of the major objectives by making pairwise comparisons.
- Have the decision maker(s) derive priorities for the relative importance of the sub-objectives by making pairwise comparisons.
- Improve and/or replace the formulas. It is difficult, if not impossible, to measure preferences with 'data based' formulas. Some of the formulas used in the book might be reasonable as crude approximations, but many are not. Some provide too much spread, some not enough. None include non-linearities, which might be very important. In general, formulas do not measure utility with respect to meeting objectives.
- Extract the most attractive (seven or so) alternatives from the initial ratings, and perform sensitivity analyses. Revise judgments and/or model structure as necessary. Refine alternative priorities derived from ratings with more accurate priorities derived from pairwise comparisons. The extracted set of alternatives could influence the judgments or model structure.

[58] Another problem with Savageau's approach is that many of the questions used subcriteria to generate the counts, but values for these subcriteria were not used in the scoring process. For example, one question asked if the cost of living was more important than historic homes in an area, the latter supposedly being used to determine the importance of housing. Yet the formula for housing did not consider historic homes at all.

> For example, if each of the alternatives being considered is relatively safe, then safety might not be as important. Similarly, if one assumes that only safe neighborhoods will be considered then crime data for the entire area might not be appropriate.

Incremental improvements such as these are very general and can be applied to *any* decision-making or evaluation methodology in *any* organization. However, decision-making and evaluation processes have significant impact on the power structure of most organizations. Thus, they are difficult to change. A sequence of incremental improvements is often wiser than trying to re-engineer or develop a 'perfect' evaluation process because such improvements will, one small step at a time, add competitiveness, demonstrate credibility, and make people feel more comfortable with the resulting power shifts.