Group Decision Making in Higher Education Using the Analytic Hierarchy Process
Author(s): Matthew J. Liberatore and Robert L. Nydick
Source: *Research in Higher Education*, Vol. 38, No. 5 (Oct., 1997), pp. 593-614
Published by: Springer
Stable URL: http://www.jstor.org/stable/40196250

Research in Higher Education, Vol. 38, No. 5, 1997

GROUP DECISION MAKING IN HIGHER EDUCATION USING THE ANALYTIC HIERARCHY PROCESS

Matthew J. Liberatore and Robert L. Nydick

The purpose of this paper is to illustrate how the analytic hierarchy process (AHP) can be applied to those situations in higher education where a group must evaluate a large number of alternatives. The suggested approach is illustrated using a case study that considers the evaluation of academic research papers at Villanova University. Following the discussion of this successful case study, a second example indicates how the AHP can be applied to the more complex problem of institution-wide strategic planning. These examples clearly demonstrate that the AHP approach is a versatile tool that can be applied to a wide range of important academic evaluation problems.

Limited internal resources and heightened competition for external resources require that many universities critically examine their decision-making processes. White (1987) reviewed the use of analytical support tools within the academic setting. He found that reported applications varied across the spectrum of decision-making activities and include planning, resource allocation, evaluation, budgeting, and scheduling. However, at the time of this study, only a few successful decision support systems had been developed and used.

Previously, the available decision support tools generally assumed the point of view of a single decision maker evaluating a set of alternatives with respect to one overriding criterion, such as cost or time. However, many important decision-making problems in academia are made by committees or groups where a number of distinct viewpoints emerge and a final decision is possible only after a consensus is reached. Today, flexible, easy-to-use, Windows-based software packages are available to support group-level decision making when multiple evaluation factors and viewpoints are present. One such methodology,

Matthew J. Liberatore, Department of Management, College of Commerce and Finance, 800 Lancaster Avenue, Villanova University, Villanova, PA 19085-1678. Robert L. Nydick, Department of Management, College of Commerce and Finance, 800 Lancaster Avenue, Villanova University, Villanova, PA 19085-1678.

0361-0365/97/1000-0593$12.50/0

the analytic hierarchy process (AHP), has been effectively applied in a variety of decision-making settings.

The next section critically reviews the AHP and competing methodologies. The discussion shows that the AHP can be an effective approach for supporting a variety of academic planning and evaluation problems. Examples of these problems include, but are not limited to, evaluating: candidates to invite for a campus interview; internal requests for research support; sabbatical proposals; faculty performance for promotion or compensation; applicants for admissions; requests for financial aid, scholarships, fellowships, or student awards; and textbooks for classroom usage. In addition, the AHP can also address institution-wide problems such as strategic planning and budgeting.

The application of the AHP through a case study follows. The context of this case is a committee's ranking of academic research papers published by faculty at one university. The in-depth presentation of this case illustrates a generic process that can be applied to any of the application areas mentioned above. One of these applications, namely the problem of institution-wide strategic planning, is examined in the subsequent section. Both examples clearly demonstrate that the AHP approach is a versatile decision support tool that should be of great interest to the academic community.

LITERATURE REVIEW

There are many methods and techniques that have been developed for the systematic evaluation of alternatives in various decision-making settings. This discussion is limited to methods and systems that can he applied in the presence of multiple criteria. These methods include goal programming (GP), multiattribute utility theory (MAUT), scoring models, and the analytic hierarchy process (AHP).

GP is a mathematical procedure for evaluating alternatives in the presence of limited resources (Lee, 1972). This technique is an improvement over standard mathematical programming techniques that consider only one criterion or objective. However, GP is a model and not a process, and thus provides no methods for ensuring that the goals selected adequately reflect the factors related to the group making the decision. In addition, GP can only incorporate criteria that can be measured quantitatively.

MAUT can be used to model the unique preferences of a decision-making group using utility functions that must be derived (Keeney and Raiffa, 1976). Practical applications have been limited due to the difficulties in constructing utility functions for an individual or group. Specifically, individuals must evaluate a sequence of artificially constructed lotteries to calibrate the utility functions for each criterion.

In addition, MAUT's view of rationality in decision making requires perfect consistency of judgments. However, in the real world, some inconsistency is acceptable and even natural. For example, in a sporting contest, if team A is twice as likely to beat team B, and if team B is three times as likely to beat team C, this does not necessarily imply that team A is six times as likely to beat team C. Inconsistency may result because of the way that the teams "match up" overall. The point is not to stop inconsistency from occurring but rather to measure the inconsistency and keep it within reasonable limits.

Scoring models are perhaps the oldest and most familiar class of multicriteria models. Requirements include developing a list of evaluation criteria, reaching a consensus on the weights given to each, and scoring each alternative with respect to each of the criteria. A weighted average score is then computed and used in ranking and selecting the alternatives. These scoring models are easy to use and construct, but their results are often misunderstood.

Typically, scoring models will use either a 1 to 5 ordinal or interval scale to rate an alternative with respect to a given criterion. However, multiplying numbers from either scale by criteria weights and summing over all criteria to obtain a total score leads to problems in interpreting the results. Specifically, numbers from either ordinal or interval scales cannot be multiplied to obtain meaningful results. That is, a score of 20 is not necessarily twice as good as a score of 10 if ordinal or interval scales are used. Such comparisons can only be made if absolute or ratio measurement scales are used (Saaty, 1980, 1982).

The AHP, developed by Saaty (1977, 1980, 1982), is a decision-making method for prioritizing alternatives when multiple criteria must be considered. The approach allows the decision maker to structure complex problems in the form of a hierarchy or a set of integrated levels. For example, a typical hierarchy will have at least three levels: the goal, the criteria, and the alternatives. In cases where there are a large number of alternatives, the hierarchy has at least three levels: the goal, the criteria, and the rating scale. In this situation the alternatives are rated with respect to each criterion.

The AHP can successfully address all of the limitations of GP, MAUT, and scoring models mentioned above. The primary advantage of the AHP is its use of pairwise comparisons to obtain a ratio scale of measurement. Ratio scales are a natural means of comparison among alternatives and enable us to measure both tangible and intangible factors. An AHP analysis uses pairwise comparisons to measure the impact of items on one level of the hierarchy on the next higher level.

Another important advantage of the AHP is that it allows for inconsistency in judgment. However, AHP also measures the degree to which the judgments are inconsistent and establishes an acceptable tolerance level for the degree of inconsistency.

Other advantages and the disadvantages of AHP have been extensively described and debated elsewhere. For example, a series of articles in *Management*

Science (Dyer, 1990a, 1990b; Harker and Vargas, 1990; Saaty, 1990; Winkler, 1990) addresses the comparisons of AHP with MAUT. Some researchers, primarily MAUT advocates, believe that allowing inconsistency of judgment and the possibility that the rankings of alternatives can change if new alternatives are added into the analysis (rank reversal) are disadvantages of using the AHP. Recent developments in the theory of AHP have addressed, and for all practical purposes resolved, these issues. These developments have led to new options in the most recent version of the AHP software package, called *Expert Choice for Windows, Version 9.0* (Expert Choice, Inc., 1995).

The AHP has been widely and successfully applied in a variety of decision-making environments (Golden, Wasil, and Harker, 1989; Zahedi, 1986; Vargas and Zahedi, 1993; Wasil and Golden, 1991). Academic administrative applications of the AHP include faculty evaluation (Lootsma, 1980; Saaty and Ramanujam, 1983; Trout and Tadisina, 1992; Tummala and Sanchez, 1988), university strategic planning (Saaty and Rogers, 1976), university budgeting (Arbel, 1983; Kwak and Diminnie, 1987), and MBA curriculum redesign (Hope and Sharpe, 1989). On a more personal level, the AHP has also been applied to help an individual select a doctoral program (Tadisina and Bhasin, 1989) and make a career choice (Canada et al., 1985) . These examples illustrate that the AHP can be a valuable decision-making tool within the academic environment. Based on the review of the available multicriteria methods, the AHP was used in the case study that follows.

CASE STUDY: RANKING RESEARCH PAPERS

Background

The College of Commerce and Finance at Villanova University has a research awards program that publicly recognizes the outstanding efforts of its faculty.[1] A secondary purpose is to foster an atmosphere that stimulates research productivity throughout the university. The college's research committee was assigned the responsibility of developing and administering the awards program. This committee consists of ten faculty members drawn from all departments within the college. All faculty members have the opportunity to submit to the committee any articles published during the previous academic year. Each paper is evaluated by members of the committee as well as an external advisory board drawn from the local business community. The committee then forwards their recommendations for awards to the dean, who makes the final decisions. Historically, the college has awarded two or three cash prizes annually.

Three interrelated problems complicated the process of evaluating research papers. First, there is often considerable variation in the papers' topics and methodology. Specifically, papers can range from theoretical to empirical and can encompass many different disciplines or subject areas. This variation contributed to the second problem, namely, disagreement about an appropriate set of evaluation criteria. Third, the committee was generally dissatisfied with the method used to process and combine individual evaluators' judgments. For these reasons, the committee chairperson asked the two authors to investigate alternative methods for evaluating and ranking the research papers. The findings of this investigation are summarized in the "Literature Review" section above and resulted in the recommendation that the AHP be applied to the evaluation of research papers. This recommendation was accepted by all members of the committee.

Structuring the Hierarchy

The development of an AHP-based evaluation system required that the committee first agree on a set of evaluation criteria, establish rating scales for each criterion, and then determine weights associated with the criteria and rating scales. After obtaining feedback on evaluation criteria from all committee members, a subcommittee drafted a list of five criteria and their definitions. After some modifications of their definitions, the criteria were approved by the full committee and have remained in effect during the implementation period. The criteria and their definitions are shown in Exhibit 1. These criteria have proven broad enough to address the differences in the types of papers entered into the competition.

EXHIBIT 1. Evaluation criteria and their definitions for ranking research papers.

OBJECTIVES: ARE THE OBJECTIVES OF THE RESEARCH CLEAR? IS THE INTENDED PURPOSE CLEAR?

JUSTIFICATION: IS A CLEAR RATIONALE PRESENTED FOR THE RESEARCH? IS THE RESEARCH POSITIONED IN LIGHT OF EXISTING KNOWLEDGE IN THE AREA? IS IT CLEAR HOW THE RESEARCH WILL EXTEND THE BODY OF KNOWLEDGE IN THIS AREA?

DESIGN: IS THE RESEARCH DESIGN (METHODOLOGY) APPROPRIATE FOR THE TOPIC? IS THE RESEARCH DESIGN ADEQUATE TO REACH THE INTENDED OBJECTIVES? WOULD OTHER APPROACHES HAVE BEEN BETTER?

EXECUTION-IMPLEMENTATION: HAS THE RESEARCH DESIGN BEEN ADEQUATELY IMPLEMENTED? WERE RESEARCH PROCEDURES EXECUTED IN SCIENTIFIC FASHION? ARE THERE ANY SHORTCOMINGS THAT MIGHT HAVE COMPROMISED THE RESEARCH?

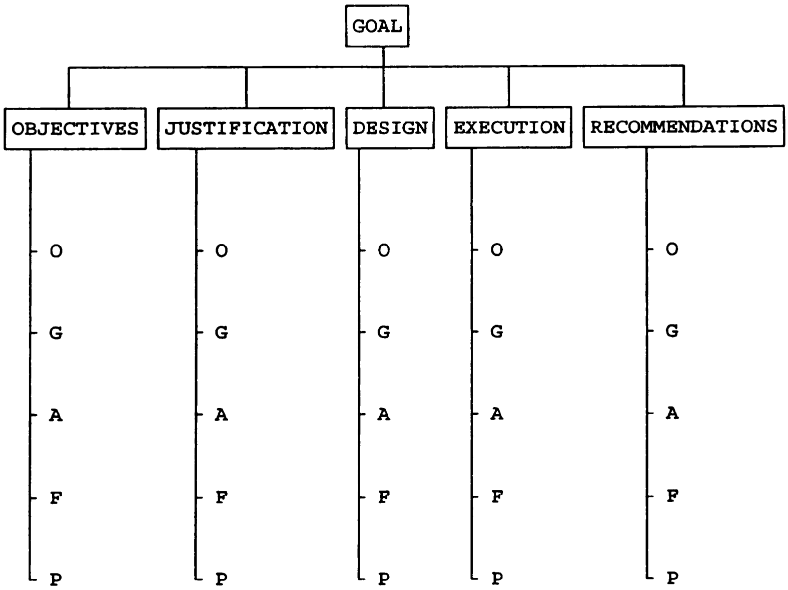
RECOMMENDATIONS AND IMPLICATIONS: DO RECOMMENDATIONS FLOW LOGICALLY FROM THE RESEARCH RESULTS? ARE FUTURE DIRECTIONS FOR RESEARCH ADEQUATELY SPECIFIED? ARE IMPLICATIONS DEVELOPED SO AS TO ADEQUATELY PLACE RESULTS IN PERSPECTIVE?

Since a large number of papers are entered into the competition each year, the committee decided to use the AHP rating scale approach. As a result, rating scales were established for each criterion. The committee decided to use the following five-point scale for all criteria: OUTSTANDING, GOOD, AVERAGE, FAIR, POOR. Each of these ratings was used by the committee members to assess the degree to which a specific paper achieved a particular criterion. Combining the five criteria with the five-point rating scale results in the hierarchy shown in Exhibit 2.

Establishing Criteria Weights

Next the committee established weights for the criteria and the rating scale categories. First, the five criteria were pairwise compared to determine their

EXHIBIT 2. AHP hierarchy for ranking research papers.



RATING SCALE
O: OUTSTANDING
G: GOOD
A: AVERAGE
F: FAIR
P: POOR

relative importance in attaining the goal. Second, the rating scale categories (OUTSTANDING, GOOD, AVERAGE, FAIR, and POOR) were pairwise compared to determine their relative importance with respect to each criterion. The hierarchy was then "synthesized" to determine the final weights associated with each rating scale category for each criterion (for example, the weight assigned to a "GOOD" OBJECTIVES rating). These pairwise comparisons were quantified using the standard one-to-nine AHP scale (Exhibit 3). This scale has been applied to a variety of decision-making problems and it has been validated in situations where standard measures (such as area and weight) already exist (Saaty, 1980).

The pairwise comparison information is represented by a pairwise compari-

EXHIBIT 3. Standard AHP 1–9 measurement scale.

Intensity of Importance	Definition	Explanation
1	Equal importance	Two activities contribute equally to the objective.
3	Moderate importance	Experience and judgment slightly favor one activity over another.
5	Strong importance	Experience and judgment strongly favor one activity over another.
7	Very strong or demonstrated importance	An activity is favored very strongly over another, its dominance demonstrated in practice.
9	Extreme importance	The evidence favoring one activity over another is of the highest possible order of affirmation.
2, 4, 6, 8	For compromise between the above values	Sometimes one needs to interpolate a compromise judgment numerically because there is no good word to describe it.
1.1–1.9	For tied activities	When elements are close and nearly indistinguishable; moderate is 1.3 and extreme is 1.9.
Reciprocals of above	If activity A has one of the above numbers assigned to it when compared with activity B, then B has the reciprocal value when compared to A	For example, if the pairwise comparison of A to B is 3.0, then the pairwise comparison of B to A is 1/3.

Source: Adapted from Saaty (1994).

son matrix. The judgments obtained from one committee member are considered first, and then the process of combining the committee's judgments is addressed. With respect to the criteria, the AHP analysis requires ten pairwise comparisons: OBJECTIVES as compared to the other four criteria, JUSTIFICATION to three criteria, DESIGN to two criteria, and EXECUTION to RECOMMENDATIONS. These judgments are entered into the upper triangular portion of the matrix.

The criteria judgments are obtained by asking questions such as the following: With respect to our goal of ranking the research papers, how much more important is the OBJECTIVES criterion than the JUSTIFICATION criterion? If an individual committee member's response is equally to moderately more important, a "2" is entered into the first row and second column in their pairwise comparison matrix. This process would be repeated for the remaining nine cases. These ten judgments are entered in the upper triangular portion of their pairwise comparison matrix (Exhibit 4). Ones are placed on the main diagonal since any item is equally important to itself. The entries in the lower triangular portion of the matrix are simply the reciprocal's of the corresponding entries in the upper triangular portion. For example, the pairwise comparison of JUSTIFICATION to OBJECTIVES is simply the reciprocal of OBJECTIVES's to JUSTIFICATION's comparison (i.e., ½ is entered in the second row and first column).

The pairwise comparison information can be used to determine the weight of each criterion. Since any matrix A of pairwise comparisons is positive and reciprocal, the weights or priorities (as they are sometimes termed in AHP) can be determined by solving the principal eigenvalue problem:

$$Aw = L_{max}\, w,$$

where w is the vector of weights and L_{max} is the maximum eigenvalue of the pairwise comparison matrix A (see Saaty, 1980).

The user need not be concerned with computing eigenvalues since there are several approaches for obtaining the weights. For example, Exhibit 5 illustrates a simple three-step procedure that provides a good approximation for the crite-

EXHIBIT 4. Criteria pairwise comparison matrix for one judge.

	OBJECTIVES	JUSTIFICATION	DESIGN	EXECUTION	RECOMMENDATIONS
OBJECTIVES	1	2	1/4	1/3	2
JUSTIFICATION	1/2	1	1/4	1/2	2
DESIGN	4	4	1	2	5
EXECUTION	3	2	1/2	1	3
RECOMMENDATIONS	1/2	1/2	1/5	1/3	1

EXHIBIT 5. Three-step approximation procedure to compute criteria weights.

1. SUM THE ELEMENTS IN EACH COLUMN OF THE ORIGINAL MATRIX.
2. DIVIDE EACH ELEMENT IN THE ORIGINAL MATRIX BY ITS COLUMN SUM. THIS RESULTS IN THE ADJUSTED MATRIX.
3. COMPUTE THE ROW AVERAGES—THESE ARE THE WEIGHTS.

A. ORIGINAL CRITERION PAIRWISE COMPARISON MATRIX

	OBJECTIVE	JUSTIFIC	DESIGN	EXECUTION	RECOMMEND
OBJECTIVE	1	2	1/4	1/3	2
JUSTIFIC	1/2	1	1/4	1/2	2
DESIGN	4	4	1	2	5
EXECUTION	3	2	1/2	1	3
RECOMMEND	1/2	1/2	1/5	1/3	1
COLUMN TOTALS	18/2	19/2	44/20	25/6	13

B. ADJUSTED RESEARCH CRITERION PAIRWISE COMPAIRSON MATRIX

	OBJECTIVE	JUSTIFIC	DESIGN	EXECUTION	RECOMMEND	WEIGHTS (ROW AVG)
OBJECTIVE	2/18*	4/19	5/44	2/25	2/13	0.1338
JUSTIFIC	1/18	2/19	5/44	3/25	2/13	0.1097
DESIGN	8/18	8/19	20/44	12/25	5/13	0.4369
EXECUTION	6/18	4/19	10/44	6/25	3/13	0.2484
RECOMMEND	1/18	1/19	4/44	2/25	1/13	0.0712
					TOTAL	1.0000

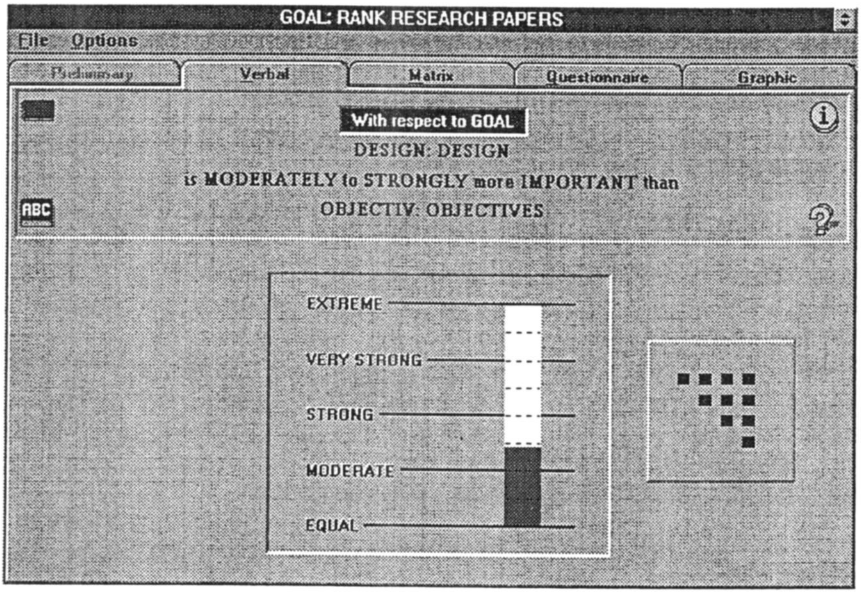
*This is obtained by dividing the RESEARCH–TEACHING pairwise comparison in the original matrix (1) by the first column total (18/2).

ria weights. This procedure can easily be programmed in any standard spreadsheet package, such as *Excel, Quattro Pro*, or *Lotus 1-2-3*.[2]

A second approach for computing the weights is to use one of several user-friendly AHP software packages, such as *Expert Choice for Windows, Version 9.0* (Expert Choice, Inc., 1995). *Expert Choice* enables the user to quickly and easily structure hierarchies, enter all necessary judgments, and automatically compute the exact weights using the eigenvalue approach. For example, after structuring the hierarchy, the user can enter judgments in several modes. Two of the more popular modes are verbal and numerical. Exhibit 6 shows how judgments are entered in the verbal mode, while Exhibit 7 shows the results of computing the weights.

The differences in weights between those given by *Expert Choice* (Exhibit 7) and the approximation procedure (Exhibit 5) are generally fairly small. The

EXHIBIT 6. Screen shot from Expert Choice showing verbal input of a judgment.



reason that *Expert Choice* is preferred is that it is easy to use and provides a full range of options for structuring, entering, editing, and analyzing hierarchies. Exhibit 7 shows that, with respect to our goal, OBJECTIVES is about 0.133/0.107 = 1.243 times more important than JUSTIFICATION when all of the judgments in the pairwise comparison matrix are considered.

Measuring Inconsistency of Judgments

An inspection of Exhibit 7 shows that the criteria matrix is somewhat inconsistent. This can be explained by considering the pairwise comparisons associated with any set of three criteria. For example, DESIGN is judged to be twice as important as EXECUTION, EXECUTION is three times as important as RECOMMENDATIONS, and DESIGN is five (not six) times as important as RECOMMENDATIONS.

As mentioned above, one important advantage of AHP is that it measures the degree to which the judgments are consistent. Inconsistency is measured in the AHP by computing a consistency ratio (CR) for each pairwise comparison matrix. The CR is defined as:

$$CR = CI/RI,$$

EXHIBIT 7. Expert Choice output after entering criteria pairwise comparisons.

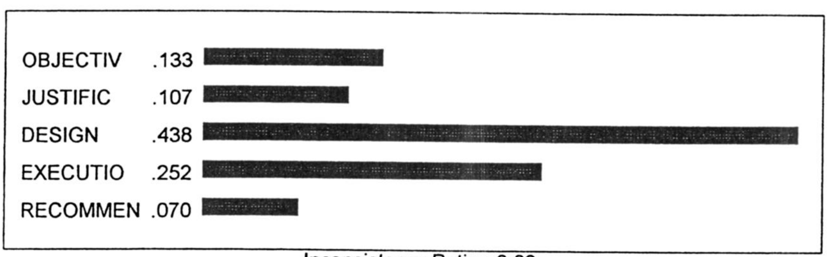
Node: 0

Compare the relative IMPORTANCE with respect to: GOAL

	JUSTIFIC	DESIGN	EXECUTIO	RECOMMEN
OBJECTIV	2.0	(4.0)	(3.0)	2.0
JUSTIFIC		(4.0)	(2.0)	2.0
DESIGN			2.0	5.0
EXECUTIO				3.0

Row element is __ times more than column element unless enclosed in ()

Abbreviation	Definition
Goal	Ranking Research Papers
OBJECTIV	OBJECTIVES
JUSTIFIC	JUSTIFICATION
DESIGN	DESIGN
EXECUTIO	EXECUTION
RECOMMEN	RECOMMENDATIONS



Inconsistency Ratio =0.03

where CI is the consistency index = $(L_{max} - n)/(n - 1)$, n is the number of items being compared, and RI is the random index, which adjusts the CI for the randomness found in an experimental set of pairwise comparison matrices (Saaty, 1980). If our pairwise comparison matrix is perfectly consistent, it can be shown that $L_{max} = n$. Therefore, the degree to which $L_{max} > n$ is a natural measure of consistency. The CR for each pairwise comparison matrix is automatically provided by *Expert Choice*.

A pairwise comparison matrix is perfectly consistent if the CR is 0.00. As long as the CR is within an order of magnitude of perfect consistency, that is, the CR is less than 0.10, it is considered to be acceptable. Exhibit 7 shows that the CR = 0.03, which is fine. If the CR is greater than 0.10, some revision of judgment may be required. *Expert Choice* also provides the user with assistance in identifying the most inconsistent judgments.

Determining the Rating Scale Weights

The next step is to determine the importance of the rating scale categories with respect to each criterion. These ratings must be pairwise compared to establish their weights. This process leads to ratio-scaled measures for the rating scale categories. The rating scale judgments are obtained by asking questions such as the following: With respect to OBJECTIVES, to what extent is an OUTSTANDING rating preferred to a GOOD rating? Nine additional comparisons are needed to complete the pairwise comparison matrix. Once again, the weights for these ratings are computed by *Expert Choice*. Since all criteria are not weighted equally, the OBJECTIVES rating weights must be adjusted by the OBJECTIVES criterion weight.

Consider the following example. Suppose the OBJECTIVES rating weights for OUTSTANDING, GOOD, AVERAGE, FAIR, and POOR turned out to be 0.400, 0.300, 0.150, 0.100, and 0.050, respectively. These weights were normalized by dividing by the highest weight (0.400) to yield 1.000, 0.750, 0.375, 0.250, and 0.125, respectively. If the OBJECTIVES criterion weight is 0.297, then the adjusted OBJECTIVES rating weights become 0.297 (1.00*0.297), 0.223, 0.111, 0.074, and 0.037, respectively.

Different sets of weights would be determined for the ratings under the other four criteria. When evaluating a paper, an individual simply selects a rating for each criterion. The weights associated with each of the five selected ratings are added to obtain a final weighted score for the research paper.

Modification for Group Decision Making

Thus far, the discussion has assumed that only one individual is providing the judgments. Since this is rarely the case in academia, modifications for group decision making are needed. Suppose each group member provides a complete set of pairwise comparisons for the criteria and the ratings. A method is required to combine these judgment sets in such a way that a consensus is obtained. Two acceptable approaches are briefly described below.

One approach for combining these judgments requires the group to resolve their differences by seeking a consensus through discussion. A second requires computing the geometric mean of each of the judgments (Aczel and Saaty, 1983). For example, suppose that there are five group members and each has provided a judgment comparing the relative importance of OBJECTIVES as compared with JUSTIFICATION. If 3, 2, 4, 3, and 4 are the five judgments, then their geometric mean is the fifth root of their product ($[(3)(2)(4)(3)(4)]^{1/5}$), which equals 3.1. The value of 3.1 is entered into the pairwise comparison matrix in the appropriate cell. This process is repeated for all judgments as needed.

The committee elected to utilize the geometric mean approach for combining the committee members' judgments. This approach was selected to minimize the time required to initialize the system. Input forms were used to elicit the necessary pairwise comparisons for criteria and rating scale weights from the various committee members. These forms were found to expedite the overall evaluation process. Examples of these forms are shown in Exhibits 8 and 9. *Expert Choice* allows the user to enter all the necessary judgments and then the software automatically computes the geometric mean for each cell.

Results

Based on some hypothetical data, Exhibit 10 summarizes the criteria and rating scale pairwise comparison matrices and associated weights. For purposes of illustration, the rating scale results shown in Exhibit 10 are used for all five evaluation criteria. Using the normalization procedure previously described, the final weights for the rating scale under each criterion can be determined. From Exhibit 10, the rating weights for OUTSTANDING, GOOD, AVERAGE, FAIR, and POOR are 0.533, 0.250, 0.124, 0.061, and 0.033, respectively. These weights were normalized by dividing by the highest weight (0.533) to yield 1.000, 0.469, 0.233, 0.114, and 0.062, respectively. Since the OBJECTIVES criterion weight is 0.106, the adjusted OBJECTIVES rating weights become 0.106

EXHIBIT 8. Sample AHP input form for eliciting judgments about criteria.

For each comparison, evaluate the relative importance of the options by placing a number next to the preferred option.

Example 1: If OBJECTIVE is **Strongly Preferred** or **Strongly More Important** than JUSTIFICATION, then:

__5__ **Objective** as compared to **Justification** _____

Example 2: If JUSTIFICATION is **Strongly Preferred** or **Strongly More Important** than OBJECTIVE, then:

_____ **Objective** as compared to **Justification** __5__

_____ OBJECTIVE as compared to JUSTIFICATION _____
_____ OBJECTIVE as compared to DESIGN _____
_____ OBJECTIVE as compared to EXECUTION _____
_____ OBJECTIVE as compared to RECOMMENDATIONS _____
_____ JUSTIFICATION as compared to DESIGN _____
_____ JUSTIFICATION as compared to EXECUTION _____
_____ JUSTIFICATION as compared to RECOMMENDATIONS _____
_____ DESIGN as compared to EXECUTION _____
_____ DESIGN as compared to RECOMMENDATIONS _____
_____ EXECUTION as compared to RECOMMENDATIONS _____

EXHIBIT 9. Sample AHP input form for eliciting judgments about rating scale.

RATING SCALE JUDGMENTS FORM FOR CRITERIA: OBJECTIVES

EXAMPLE: IF AN **OUTSTANDING** RATING IS STRONGLY PREFERRED OR STRONGLY MORE IMPORTANT THAN A **GOOD** RATING, ENTER A 5 IN THE SPACE PROVIDED.

_____ **OUTSTANDING** RATING AS COMPARED TO A **GOOD** RATING
_____ **OUTSTANDING** RATING AS COMPARED TO A **AVERAGE** RATING
_____ **OUTSTANDING** RATING AS COMPARED TO A **FAIR** RATING
_____ **OUTSTANDING** RATING AS COMPARED TO A **POOR** RATING
_____ **GOOD** RATING AS COMPARED TO A **AVERAGE** RATING
_____ **GOOD** RATING AS COMPARED TO A **FAIR** RATING
_____ **GOOD** RATING AS COMPARED TO A **POOR** RATING
_____ **AVERAGE** RATING AS COMPARED TO A **FAIR** RATING
_____ **AVERAGE** RATING AS COMPARED TO A **POOR** RATING
_____ **FAIR** RATING AS COMPARED TO A **POOR** RATING

(1.00*0.106), 0.050, 0.025, 0.012, and 0.006, respectively. Executing the ratings command in *Expert Choice* performs all of these necessary computations. The committee reviewed the consensus weights and was satisfied with the results.

Typically, 10 to 15 papers have been entered in the competition annually and are evaluated by approximately 14 judges (both external and internal). To illustrate the process, some hypothetical data for a simplified example are presented. Assume that fifteen papers are to be evaluated by five judges.

A standard scoring sheet is used to capture the judgments of each committee member for each paper. The committee members are aware of the rating scale weights for each criterion but *not* the criteria weights themselves in order to minimize potential biases. The ratings from each committee member's scoring sheet are entered into *Expert Choice* to compute the overall scores for each paper. Exhibit 11 shows a sample *Expert Choice* output for one judge.

The final step requires consolidating all of the evaluators' judgments into a summary spreadsheet. An example of a typical summary sheet for five judges is shown as Exhibit 12. Because the scores are ratio-scaled, the committee members can directly compare the final paper scores and use statistical summaries to accelerate the process of reaching a consensus. The median rankings were used as a starting point for establishing the final rankings of the papers. If the median scores of closely ranked papers are not substantially different, then the additional information presented in Exhibit 12 can be used to reach a consensus on the final rankings. An important point to remember is that the results of the AHP analysis should be used as an *aid* in the decision-making process. Expert judgment is still required in interpreting the rankings and reaching a decision. A key benefit of the AHP is that it provides a framework for structuring and analyzing a decision.

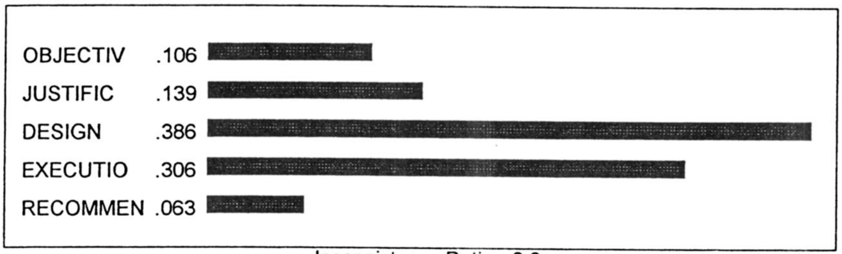
EXHIBIT 10. Abbreviated Expert Choice outputs for group-level criteria and rating scale pairwise comparisons.

Node: 0

Compare the relative IMPORTANCE with respect to: GOAL

	JUSTIFIC	DESIGN	EXECUTIO	RECOMMEN
OBJECTIV	(1.2)	(3.8)	(3.1)	1.7
JUSTIFIC		(2.4)	(2.2)	2.1
DESIGN			1.4	6.1
EXECUTIO				5.2

Row element is __ times more than column element unless enclosed in ()



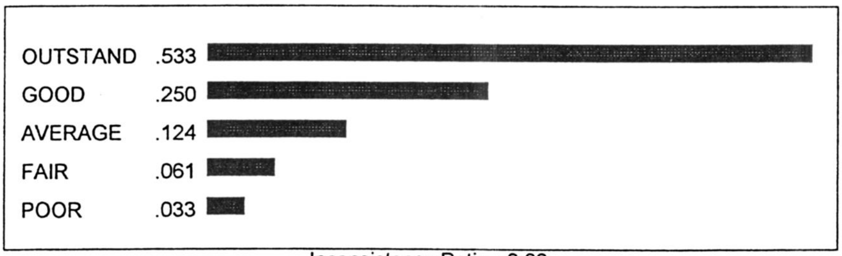
Inconsistency Ratio =0.0

Node: 10000

Compare the relative PREFERENCE with respect to: OBJECTIV < GOAL

	GOOD	AVERAGE	FAIR	POOR
OUTSTAND	3.5	5.4	7.4	9.0
GOOD		3.0	5.1	7.1
AVERAGE			3.0	4.9
FAIR				2.9

Row element is __ times more than column element unless enclosed in ()



Inconsistency Ratio =0.06

Implementation Experience

Prior to implementing the modified AHP approach, the ten committee members met as many as five times with most sessions lasting three hours or more. These sessions were required to settle upon the format of the evaluation process, discuss the merits of the papers, and achieve a consensus on the final

EXHIBIT 11. Expert Choice output summarizing results of one judge's ratings.

A:\RESEARCH

	Alternatives	TOTAL	OBJECTIV .1058	JUSTIFIC .1387	DESIGN .3865	EXECUTIO .3064	RECOMMEN .0626
1	PAPER11	0.797	AVERAGE	GOOD	OUTSTAND	OUTSTAND	AVERAGE
2	PAPER5	0.708	GOOD	GOOD	OUTSTAND	GOOD	OUTSTAND
3	PAPER9	0.640	AVERAGE	GOOD	GOOD	OUTSTAND	OUTSTAND
4	PAPER12	0.419	OUTSTAND	OUTSTAND	AVERAGE	AVERAGE	AVERAGE
5	PAPER13	0.398	GOOD	FAIR	GOOD	GOOD	FAIR
6	PAPER14	0.369	FAIR	OUTSTAND	GOOD	FAIR	POOR
7	PAPER2	0.338	OUTSTAND	AVERAGE	FAIR	GOOD	AVERAGE
8	PAPER3	0.310	GOOD	FAIR	GOOD	FAIR	GOOD
9	PAPER4	0.288	FAIR	AVERAGE	GOOD	FAIR	GOOD
10	PAPER7	0.278	AVERAGE	AVERAGE	GOOD	FAIR	FAIR
11	PAPER6	0.267	FAIR	GOOD	FAIR	GOOD	POOR
12	PAPER1	0.239	AVERAGE	AVERAGE	POOR	GOOD	AVERAGE
13	PAPER10	0.155	GOOD	FAIR	FAIR	FAIR	AVERAGE
14	PAPER8	0.150	POOR	GOOD	FAIR	FAIR	POOR
15	PAPER15	0.148	POOR	FAIR	AVERAGE	FAIR	POOR
16							

rankings. The application of the AHP system only required two meetings lasting a total of about three hours to reach complete consensus on the rankings of the top three choices. Using the new system resulted in an annual savings of approximately 120 hours of faculty members' time per year. A significant portion of this time savings resulted from the elimination of the need to "reinitialize" the evaluation system each year.

In addition to the time savings, there were other important benefits attributable to the AHP system. Most importantly, the AHP provided added structure and consistency to the evaluation process. Using the new system, it was evident how each judge's scores were determined and how they related to the final rankings of the papers. These benefits led to the belief that the evaluation process was less biased and less "political" than in the past. In addition, since the evaluation process remained essentially unchanged during a four-year period, consistency in evaluation was maintained across years as well. Based on positive feedback from the faculty, the committee concluded that the AHP offered substantial benefits over the previous system.

As would be expected, there were some start-up costs associated with the development of the AHP system. During the initial year, each committee member spent two hours learning about the AHP system and providing the necessary criteria definitions and related judgments. In addition, the authors spent about ten hours collecting and entering the judgments into *Expert Choice*. After the first year, using the AHP system required about five hours of the authors' time annually. Since the *Expert Choice* software was already available on the college's local area network (LAN), no additional hardware or software costs were incurred.

The criteria and rating scales were reviewed each year and were revised after

EXHIBIT 12. Sample summary sheet of judge's ratings of research papers.*

	JUDGE 1		2		3		4		5						# OF	# OF	# OF
PAPER	SCORE	RANK	SCORE	RANK	SCORE	RANK	SCORE	RANK	SCORE	RANK	MEDIAN	MEDIAN RANK	MEAN	MEAN RANK	1ST RANK	2ND RANK	3RD RANK
1	0.239	12	0.285	12	0.346	10	0.144	12	0.402	8	0.239	12	0.283	12	0	0	0
2	0.338	7	0.332	6	0.442	5	0.238	10	0.436	7	0.338	7	0.357	6	0	0	0
3	0.310	8	0.292	11	0.299	12	0.242	8	0.292	12	0.292	9	0.287	11	0	0	0
4	0.288	9	0.273	13	0.362	9	0.299	5	0.255	14	0.288	10	0.295	10	0	0	0
5	0.708	2	0.642	3	0.441	6	0.551	1	0.601	3	0.601	2	0.589	2	1	1	2
6	0.267	11	0.314	9	0.293	13	0.197	11	0.511	6	0.293	8	0.316	8	0	0	0
7	0.278	10	0.304	10	0.394	8	0.243	7	0.265	13	0.278	11	0.297	9	0	0	0
8	0.150	14	0.221	15	0.301	11	0.052	15	0.322	10	0.221	14	0.209	14	0	0	0
9	0.640	3	0.702	2	0.551	3	0.485	2	0.544	5	0.551	3	0.584	3	0	2	2
10	0.155	13	0.321	8	0.232	15	0.096	13	0.307	11	0.232	13	0.222	13	0	0	0
11	0.797	1	0.745	1	0.714	1	0.398	3	0.560	4	0.714	1	0.643	1	3	0	1
12	0.419	4	0.423	4	0.711	2	0.365	4	0.740	1	0.423	4	0.532	4	1	1	0
13	0.398	5	0.414	5	0.443	4	0.245	6	0.602	2	0.414	5	0.420	5	0	1	0
14	0.369	6	0.324	7	0.420	7	0.239		0.365	9	0.365	6	0.343	7	0	0	0
15	0.148	15	0.222	14	0.245	14	0.092	14	0.219	15	0.219	15	0.185	15	0	0	0

*The scores shown for Judge 1 were obtained from the spreadsheet given as Exhibit 11. The paper scoring process would then be repeated for the other judges with the results transferred to a summary sheet similar to the one shown above. (The supporting spreadsheets for Judges 2 through 5 are not shown.)

three years. This required the committee to spend an additional two hours per person to support the revision process. It is important to note that this revision was *not* precipitated by a change in the college's research goals, but rather by a considerable change in the composition of the research committee's membership. (The latter is part of a standard college procedure for cycling faculty through a given committee.) This review provided the new committee members with an opportunity to modify the criteria, the rating scales, and the weights associated with each. As a result, the committee decided to maintain the same set of criteria and their definitions, while making some minor modifications to the weights. More importantly, this revision process enabled all of the committee members to "buy into" the evaluation process.

Based on the favorable analysis of benefits and costs, the committee and the faculty judged that the AHP system should continue to be used. As a result, the AHP system has been used successfully for four years. The dean has accepted the committee's recommendations each year.

Drawing upon the experiences gained from the case study, the next section describes how the AHP can be applied to the more complex problem of institution-wide strategic planning.

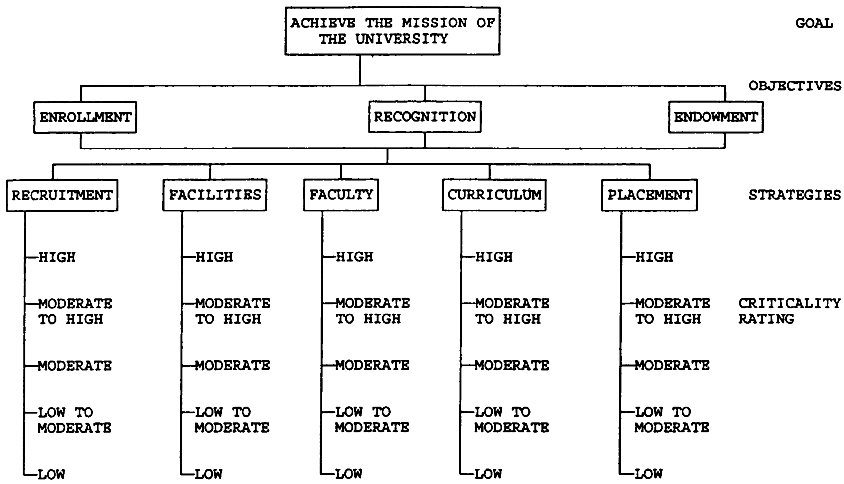
A SECOND EXAMPLE: INSTITUTION-WIDE STRATEGIC PLANNING

Developing a Strategic Planning Hierarchy

Consider a university that wishes to formalize its strategic planning process and reach a consensus on the action plans that will be pursued over a five-year planning horizon. Our suggested approach follows a well-known format for planning that first requires developing the mission or charter of the organization. The internal and external environments are then analyzed, often using a SWOT (strengths, weaknesses, opportunities, and threats) analysis. Next, objectives that are time based and measurable are specified. Lastly, strategies and action plans are developed to achieve these objectives. This planning approach is usually called mission, objectives, and strategies (MOS).

In applying the MOS approach, action plans usually are rated according to their criticality in achieving specific strategies. Since strategies contribute toward the achievement of objectives, and in turn the mission of the university, a clear and formal linkage can be established between the process of prioritizing action plans and the mission. The AHP can be used to quantify this linkage and hence support the strategic planning process. An example of an MOS hierarchy for this university is given as Exhibit 13, while a list of representative action plans for each strategy is given as Exhibit 14.

EXHIBIT 13. Sample AHP hierarchy for university strategic planning.



Processing the Hierarchy

The AHP process begins by pairwise comparing the objectives according to their importance in achieving the mission. For example, the three objectives shown in Exhibit 13 might be stated as follows:

ENROLLMENT: Increase full-time student enrollment by 3% per year over five years

RECOGNITION: Become listed in *U.S. News*'s top 50 national universities within five years

ENDOWMENT: Increase endowment by $500 million over five years

One pairwise comparison question might be phrased: "With respect to achieving the mission of the university, how much more important is the ENROLLMENT objective than the RECOGNITION objective?" With three objectives, only three pairwise comparisons are needed.

The process continues by pairwise comparing the strategies that have been identified with respect to each objective. With five strategies, this requires ten pairwise comparisons for each of the three objectives. An important benefit of this process is that the individuals involved must clearly think through the strength of the relationships between strategies and objectives. Combining the weights of the objectives with respect to the mission with the weights of the

EXHIBIT 14. Sample action plans to help achieve each strategy.

RECRUITMENT
- INCREASE VISITS TO U.S. HIGH SCHOOLS BY 20%
- VISIT SELECTED INTERNATIONAL HIGH SCHOOLS IN LATIN AMERICA
- INCREASE THE NUMBER AND SCOPE OF ON-CAMPUS RECRUITING PROGRAMS

FACILITIES
- CONSTRUCT A NEW CONVOCATION CENTER
- BUILD A STATE-OF-THE-ART STUDENT CENTER
- UPGRADE COMPUTER FACILITIES TO A MINIMUM PENTIUM CONFIGURATION
- BUILD APARTMENT-STYLE HOUSING FOR STUDENTS
- BUILD MULTIMEDIA INSTRUCTIONAL CENTER

FACULTY
- IMPLEMENT A FORMAL FACULTY DEVELOPMENT PROGRAM
- IMPLEMENT A MENTORING PROGRAM FOR NEW FACULTY
- ESTABLISH A SERIES OF WORKSHOPS TO HELP IMPROVE TEACHING EFFECTIVENESS

CURRICULUM
- DEVELOP A CORE HUMANITIES CURRICULUM
- DEVELOP A FIVE-YEAR ACCOUNTING PROGRAM
- INSTITUTE A UNIVERSITY COMPUTER LITERACY REQUIREMENT
- CREATE A JOINT DEGREE PROGRAM WITH A FOREIGN UNIVERSITY

PLACEMENT
- INCREASE STUDENT PLACEMENT TO 90% WITHIN THREE MONTHS OF GRADUATION
- INCREASE THE NUMBER OF ON-CAMPUS RECRUITERS BY 15%
- INCREASE PLACEMENT IN MEDICAL/LAW SCHOOLS BY 10%
- INSTITUTE ON-LINE PLACEMENT SERVICE USING THE WORLD WIDE WEB

strategies with respect to each objective, we obtain the weights of the strategies with respect to the mission.

The next step determines the importance of the criticality rating scale categories with respect to each strategy. These ratings must be pairwise compared to establish their weights. For example, suppose that the normalized rating scale weights for the CURRICULUM strategy are 1.000, 0.714, 0.571, 0.429, and 0.143. (As described in the case study, the weights for the ratings may differ by strategy.) If the CURRICULUM strategy weight is 0.223, then the adjusted CURRICULUM rating scale weights would be 0.223, 0.159, 0.127, 0.096, and 0.032, respectively.

When evaluating a particular action plan, an individual simply selects the appropriate criticality rating for the supported strategy. For example, the action plan to implement a computer literacy requirement might be rated as MODERATE

TO HIGH in achieving the CURRICULUM strategy. If an action plan supports more than one strategy, the action plan's total score is simply the sum of the weights of the associated criticality ratings.

Ranking the action plans in descending order of their total score is one method for deciding which action plans to pursue. However, this approach does not take into consideration the costs of accomplishing a specific action plan. A simple resolution of this problem requires ranking the action plans in descending order of their benefit/cost ratio. The benefit/cost ratio can be computed by dividing the action plan's total score by its cost. Action plans are supported until budget is no longer available.

CONCLUSION

This paper discussed the application of the AHP approach to support important group decision-making processes within the academic environment. The first example addressed a case study concerning the ranking of academic research papers at Villanova University. The second described the application of the AHP to the more complex problem of institution-wide strategic planning. These examples clearly demonstrate that the AHP approach is a versatile tool that can be applied to a wide range of important academic evaluation problems.

NOTES

1. See Liberatore, Nydick, and Sanchez (1992) for a brief discussion of this problem.
2. A macro-driven *Lotus 1-2-3* spreadsheet has been prepared by the authors to perform the AHP computations for problems having up to three criteria and four alternatives. Contact the authors if you wish to receive a copy.

REFERENCES

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

Arbel, A. (1983). A university budget problem: A priority-based approach. *Socio-Economic Planning Sciences* 17(4): 181–189.

Canada, J. R., Frazelle, E. H., Koger, R. K., and MacCormac, E. (1985). How to make a career choice: The use of the analytic hierarchy process. *Industrial Management* 27(5): 16–22.

Dyer, J. S. (1990a). Remarks on the analytic hierarchy process. *Management Science* 36(3): 249–258.

Dyer, J. S. (1990b). A clarification of "remarks on the Analytic Hierarchy Process." *Management Science* 36(3): 274–275.

Expert Choice, Inc. (1995). *Expert Choice for Windows, Version 9.0*, Pittsburgh, PA.

Golden, B. L., Wasil, E. A., and Harker, P. T. (eds.) (1989). *The Analytic Hierarchy Process*, New York, Springler-Verlag.

Harker, P. T., and Vargas, L. G. (1990). Reply to "Remarks on the Analytic Hierarchy Process" by J. S. Dyer. *Management Science* 36(3): 269–273.

Hope, R. P., and Sharpe, J. A. (1989). The use of two planning decision support systems in combination for the redesign of an MBA information technology programme. *Computers and Operations Research* 16(4): 325–332.

Keeney, R. L., and Raiffa, H. (1976). *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*. New York: Wiley.

Kwak, N. K., and Diminnie, C. B. (1987). A goal programming model for allocating operating budgets of academic units. *Socio-Economic Planning Sciences* 21(5): 333–339.

Lee, S. M. (1972). *Goal Programming for Decision Analysis*. Philadelphia: Auerbach.

Liberatore, M. J., Nydick, R. L., and Sanchez, P. M. (1992). The evaluation of research papers (or how to get an academic committee to agree on something). *Interfaces* 22(2): 92–100.

Lootsma, F. A. (1980). Saaty's priority theory and the nomination of a senior professor in operations research. *European Journal of Operational Research* 4(6): 380–388.

Saaty, T. L. (1977). A scaling method for priorities in hierarchical structures. *Journal of Mathematical Psychology* 15(3): 234–281.

Saaty, T. L. (1980). *The Analytic Hierarchy Process*. New York: McGraw-Hill.

Saaty, T. L. (1982). *Decision Making for Leaders*. Belmont, CA: Lifetime Learning Publications, Division of Wadsworth.

Saaty, T. L. (1990). An exposition of the AHP in reply to the paper "Remarks on the Analytic Hierarchy Process." *Management Science* 36(3): 259–268.

Saaty, T. L. (1994). How to make a decision: The analytic hierarchy process. *Interfaces* 24(6): 19–43.

Saaty, T. L., and, Ramanujam, V. (1983). An objective approach to faculty promotion and tenure by the analytic hierarchy process. *Research in Higher Education* 18: 311–331.

Saaty, T. L., and Rogers, L. R. (1976). Higher education in the United States (1985–2000): Scenario construction using a hierarchical framework with eigenvector weighting. *Socio-Economic Planning Sciences* 10: 251–263.

Tadisina, S. K., and Bhasin, V. (1989). Doctoral program selection using pairwise comparisons. *Research in Higher Education* 30(4): 403–418.

Trout, M. D., and Tadisina, S. K. (1992). The analytic hierarchy process as a model base for a merit salary recommendation system. *Mathematical and Computer Modelling* 11: 276–281.

Tummala, V. M. R., and Sanchez, P. P. (1988). Evaluating faculty merit awards by analytic hierarchy process. *Modelling, Simulation and Control C: Environmental, Biomedical, Human and Social Systems* 11(4): 1–13.

Vargas, L. G., and Zahedi, F. (eds.) (1993). Special issue on the analytic hierarchy process. *Mathematical and Computer Modelling* 17(4–5).

Wasil, E. A., and Golden, B. L. (eds.) (1991). Public sector applications of the analytic hierarchy process. *Socio-Economic Planning Sciences* 25(2).

White, G. P. (1987). A survey of recent management science applications in higher education. *Interfaces* 17: 97–108.

Winkler, R. L. (1990). Decision modeling and rational choice, AHP and utility theory. *Management Science* 36(3): 247–248.

Zahedi, F. (1986). The analytic hierarchy process—A survey of the method and its applications. *Interfaces* 16(4) 96–108.

Received June 5, 1996.