ISSN: 2521-0874 (Print)
ISSN: 2521-0505 (Online)

Contents List available at RAZI Publishing

Acta Informatica Malaysia(AIM)

Journal Homepage: http://www.razipublishing.com/journals/acta-informatica-malaysia

https://doi.org/10.26480/aim.01.2017.01.06

# Teaching Performance Evaluation Framework: An Analytic Hierarchy Process Approach

**Azizul Azhar Ramli, Shahreen Kasim, Mohd. Farhan Md. Fuzzee, Hairulnizam Mahdin**

Faculty of Computer Science and Information Technology Universiti Tun Hussein Onn Malaysia Parit Raja, 86400 Batu Pahat, Johor, Malaysia
{azizulr, shahreen, farhan, aizi}@uthm.edu.my



**ARTICLE DETAILS**

*Article history:*

Received 22 January 2017
Accepted 03 February 2017
Available online 05 February 2017

*Keywords:*

Teaching Performance Evaluation Process, Multi Criteria Decision Making, Analytic Hierarchical Process

**ABSTRACT**

Teaching performance evaluation process in higher learning institutions (HLI) is of the most importance in building a higher education policy which purposes for teaching improvement and administrative decision making. Several issues have been raised such as 1) instrument characteristic used to detect values is not sufficient to handle subjective evaluation criteria's and, 2) the most techniques used in handling survey data is statistical method. In conventional practice, ordinal Likert scales used for marking schema and statistical techniques are implemented for analyzing the data. In order to manage the issues, this research presents a multi-criteria decision making methodology, called analytic hierarchical process (AHP) to evaluate HLI instructors teaching performance. As a well known decision making approach, AHP can systematically cope with highly subjective problems in social system. An AHP implementation framework for teaching performance evaluation will use the possible internal factors; preparation, organization, delivery and effectiveness that are evaluated as teaching evaluation weight. The future work is also suggested, this evaluation system should be implemented in real time situations. It becomes the best platform to evaluate teaching performance and categorize instructors into appropriate bands.

## 1. Introduction

The implementation of performance evaluation is a necessary and beneficial process to provide annual feedback to staff members about job effectiveness and career guidance. The performance review is intended to be a fair and balanced assessment of an employee's performance. Currently, there are a lot of instruments that can be used in order to assist supervisors and department heads in conducting performance reviews process in higher learning institutions (HLI) world. It is important to implement teaching performance evaluation process especially in academia. Teaching and learning are a core business and continuous improvement process must be effectively done in HLI. Otherwise those HLI will not meet their standard. In addition, HLI instructors are a back bone of this education organization. Nowadays, the evaluation criteria of teaching performance of HLI instructors usually consists of several components and parameters, each involving a number of judgments often based on imprecise data where most used instruments are an optical mark reader (OMR) answer sheets to select ordinal Likert scale answers.

In addition, statistical methods are used for aggregating information from these assessment parameters. These methods are accepted especially in many HLI around the world although it is argued that this statistical technique does not necessarily offer the best way to evaluate human knowledge and skills. Hence, this study has the main objective of proposing a mechanism to evaluate the HLI instructors teaching performance evaluation. The proposed mechanism can often reflects the way people think and make judgments.

With the ability and features of an MCDM approach especially AHP, in decision making process, the quality of the teaching performance evaluation will result, more reliable and validate one with highest confidences.

The rest of the paper is organized as follows. In Section 2, some literature is reviewed about teaching performance evaluation process including its guideline, MCDM theory and AHP approach and also related research efforts to this research topic. Section 3 presents a suggested design model and methodology for teaching performance evaluation based on MCDM method. The final section, Section 4 presents some concluding remarks.

## 2. Literature Review

### A. Teaching Performance Evaluation Process

Teaching performance evaluations are often designed to serve two purposes; to measure instructor competence and to foster professional development and growth. A teaching evaluation system should give instructors useful feedback on classroom needs, the opportunity to learn new teaching techniques and counsel from principals and other instructors on how to make changes in their classrooms. Evaluators are only mostly students. Specific procedures and standards must be first set to achieve these goals evaluations. (Boyd, 1989).

Although, there are several of practices are currently applied for conducting the teaching evaluation process and with those mechanisms, there are a lot of controversies especially on the resulted reliability and validity.
Furthermore, evaluation of teaching in HLI involves collecting evidence, from various stakeholders (mainly students), for the purpose of improving the effectiveness of the teaching-learning implementation process. A successful evaluation generates valid and reliable outcomes that indicate directions and actions to improve. This is a main objective of an evaluation process. Fig. 1 shows an example of an OMR form for teaching performance evaluation process with Likert scale answers. This standard evaluation form is categorized as close-ended type of questionnaires. Beside that, there are certain HLI still used open-ended questions which are including highly subjective form of evaluation criteria and need a pedagogical expertise to analyze and interpret the conclusions results.

Based on stated mechanism above, OMR form notation (multiple-choice questions where the students indicate their answers by marking on the form which can then be automatically scanned and scored) are widely used because it is recognized as an efficient and cost effective data entry method for the large scale processing of category data. From its origins in the purposes of examinations, it is now widely used in other large-scale evaluations and exercises. Traditionally, questionnaires are evaluated using copies of paper questionnaires and/or OMR answer sheets. An optical mark scanner will then be used to read the marks made on the OMR forms. The process will produce an electronic data file containing responses to the questionnaires. The data file can then be analyzed using software programs such as Microsoft Excel, SPSS or SAS. With the availability of web technology, there is now the option of collecting evaluation data online. Recently, a number of IHL use Internet facilities for teaching performance evaluation purposes which is more convenient and effective to generate the final results.

The conventional method recently used on teaching evaluation process especially in HLI is a statistical approach which is applied mean or average score calculation. This calculation method is synonym with the usage of an OMR form which is most often used during an evaluation process. On the other hand, there are several methods that are also applied to

evaluation data such as statistical method includes studentized range or coefficient of variation, reliability analysis and so on.

HLI Instructors Teaching Performance Evaluation Guideline

Teaching performance evaluation or assessment provides instructors with starting-points for their teaching and with advices on how to gather feedback on their teaching practice and effectiveness as a part of a systematic program of teaching development.

All evaluation techniques contain implicit judgments about the characteristics that constitute quality of teaching. These judgments should be made explicit and indeed should become part of the evaluation process itself in a manner which recognizes instructors' privileges to be evaluated within the context of their own teaching philosophies and goals (Geis, 1977). The criteria for evaluating teaching performance vary among disciplines and within disciplines and should take into consideration the level of the course, the instructor's objective and style, and the teaching methodology employed. Nonetheless, the primary criterion must be to improve student learning. Concretely, indicators of quality of teaching in IHL can include the following.

i. Effective choice of materials,
ii. Organization of subject matter and course,
iii. Effective communication skills,
iv. Knowledge for the subject matter and
v. Availability and responsiveness to student concerns and opinions.

| Evaluation Parameters/Scales | Strongly agree | Agree | Undecided | Disagree | Strongly disagree |
|---|---|---|---|---|---|
| Discussing teaching activities plan and assessment clearly. | O | O | O | O | O |
| Exhibited knowledge of the subject. | O | O | O | O | O |
| Teaching material and resource were helpful. | O | O | O | O | O |
| Teaching activities provided relevant learning experiences. | O | O | O | O | O |
| Gave clear explanations and provided extra teaching materials. | O | O | O | O | O |
| Using various and effective teaching techniques. | O | O | O | O | O |
| Showed enthusiasm and concern for encouraging student learning. | O | O | O | O | O |
| Guided to achieve the understanding of subject matter. | O | O | O | O | O |

**Fig.1 A section of a typical closedquestionnaire used for teaching evaluation process**

In addition, the value added for teaching quality can be enhanced by providing the following criteria.
i. To establish a positive learning environment,
ii. To motivate student engagement,
iii. To provides appropriate challenges,
iv. To be responsive to students learning needs and
v. To be fair in evaluating their learning.

Regarding the teaching assessment methods, there are two general categories include formative and summative purposes. The formative assessment of teaching performance can be carried out at many points e.g. during instructional period either in formal classroom or virtual environment. The main purpose of the assessment is for instructors to find out the certain part that can be improved such as teaching style and student's evaluation procedure. The assessment is initiated by the instructors and solicited from many sources such as self and students using a variety of instruments including survey or online forms.

On the other hand, the summative assessment is usually conducted at the end of a particular course or at the specific points in instructor's career. The purpose is to form a judgment about the effectiveness of a courses and/or instructors. The judgment may be used for tenure and promotion decision to reward success in the form of teaching awards or merit pay, or to enable departments to make informed decisions about changes to individual courses, the curriculum or teaching assignments.
At most IHL, the summative evaluation is implemented and it includes the results of teaching evaluations regularly scheduled at the end of academic terms. Furthermore, the famous evaluation instrument for this purpose is OMR form and statistical calculation method.

**B.Multiple Criteria Decision Analysis (MCDM)**
MCDM techniques are often selected arbitrarily: sometimes an analyst is already familiar with a procedure, other times a technique is developed in an ad hoc way of the following; it also happens that a technique is chosen simply because its software is available. In fact, researchers highlight that, each situation or problem demands specific MCDM technique. The impact of the choice of a technique on actual decisions is also well known, as well as the consequences of poor decisions.

Basically, there are three generic types of solutions for MCDM problems. Table 1 shows the summary of particular types of MCDM method and this paper focus on implementation ranking solution of MCDM method.

Table 1 Solutions by MCDM Method

| | | |
|---|---|---|
| 1. | Selection | Given a set $A$ of alternatives (also called candidates), the selection task involves finding a subset $A'$ of $A$ composed as small a number of alternatives as possible, judge by decision makers to be the most satisfying. |
| 2. | Sorting | The sorting operation (also called classification) consists of assigning each alternative from $A$ to one of the number of predefined categories. The assignment should be based on the intrinsic measure of criterion for an alternative and not on its comparison with other alternatives from $A$. However in practice assignment is often based on the relative differences of alternatives on the given criterion. |
| 3. | Ranking | The ranking operation involves establishing a preference ranking on the set of alternative $A$. The preference ranking represents a priority list of the alternatives. |

*Analytic Hierarchy Process (AHP) Approach*

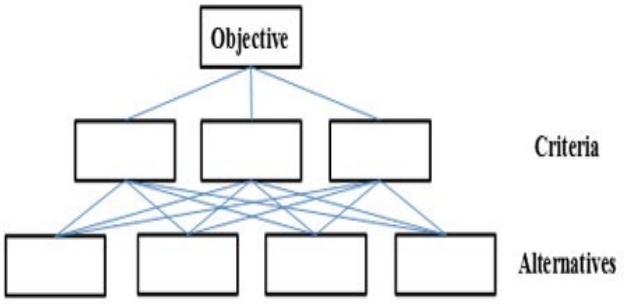


Fig. 2 Hierarchical structure of decision making problem

An alternative approach that categorized as a family of MCDM problem is an AHP approach. Regarding this popular approach, there is a long period of debate, in fact on the effective value of the AHP approach. Perez (1995) proved that the AHP approach is based on a firm theoretical foundation.

The AHP has widespread use due to its flexibility, easy to use and Kirkwood (1997) mentioned that AHP can be even implemented in spreadsheet environment. However, ambiguity in relative importance, inconsistent judgments by decision makers and the use of 1 to 9 scales can be thought as the disadvantages of this approach. The ratio scale makes sense when dealing with something like distance or areas which are natural ratio scales but not when dealing with comfort, image or quality of life, for which no clear reference levels exists. Furthermore, for large problems, too many pair-wise comparisons must be performed (Malczewski, 1999).

Discuss below the general steps of AHP approach will be discussed below. Steps are as shown below the following procedures.

In addition, simulation result will yield a set of consistency ratios for the corresponding matrices as stated in Step 6. The distribution of the consistency ratio should be studied and corresponding changes in the pair-wise comparison matrices should be made (Saaty, 1988).

***Step 1:*** Define the actual decision making problem and determine its goal.
***Step 2:*** Structure the hierarchy from the top (the objectives from decision maker's viewpoint) through the intermediate levels (criteria on which subsequent levels depend) to the lowest level which usually contains the list of alternatives.
***Step 3:*** Construct a set of pair-wise comparison matrices (size n x n) for each of the lower levels with one matrix for each element in the level immediately above by using the relative scale measurement shown in Table 2. The pair-wise comparisons are done in term of which element dominates the other.

Table 2 Pair-Wise Comparison Scale for AHP Preference

| Numerical Rating | Verbal Judgment of Preference |
|---|---|
| 9 | Extremely preferred |
| 8 | Very strongly to extremely |
| 7 | Very strongly preferred |
| 6 | Strongly to very strongly |
| 5 | Strongly preferred |
| 4 | Moderately to strongly |
| 3 | Moderately preferred |
| 2 | Equally to moderately |
| 1 | Equally preferred |

***Step 4:*** There are *n (n-1)* judgments required to develop the set of matrices in Step 3. Reciprocals are automatically assigned in each pair-wise comparison.
***Step 5:*** Hierarchical synthesis is now used to weigh the eigenvectors by the weights of the criteria and the sum is taken over all weighed eigenvector entries corresponding to those in the next lower level of the hierarchy.
***Step 6:*** Having made all the pair-wise comparisons, the consistency is determined by using the eigenvalue, $\lambda_{max}$, to calculate the consistency index, CI as follows

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

Where *n* is the matrix size. Judge consistency ratio (CR) of CI with the appropriate value in Table 3. The CR is acceptable, if it does not exceed 0.10. If it is more, the judgment matrix is inconsistent. To obtain a consistent matrix, judgment should be reviewed and improved.

Table 3 Average Random Consistency (RI)

| Size of Matrix | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Random Consistency | 0 | 0 | 0.58 | 0.09 | 1.12 | 1.24 | 1.32 | 1.41 | 1.45 | 1.49 |

***Step 7:*** Steps 3-6 are performed for all levels in the hierarchy.

**C. Related Research Efforts**

These studies are carried out to improve the statistical results of a method used in a real teaching evaluation process. This traditional analysis or synonym with Likert scale was compared with selected MCDM approach, called AHP. In general Likert scale is a method of ascribing quantitative value to qualitative data, to make it amenable to statistical analysis. Used mainly in training course evaluations and market surveys, Likert scale usually have five potential choices such as strongly agree, agree, neutral, disagree and strongly disagree, but sometimes go up to ten or more. A numerical value is assigned to each potential choice and a mean figure for all the responses is computed at the end of the evaluation or survey. The final average score represents the overall level of accomplishment or attitude toward the subject matter.

The statistical technique most frequently used is descriptive techniques which include

i. Summaries using a median or a mode. The mode is probably the most suitable for easy interpretation.
ii. Expression variability in terms of the range or inter quartile range.
iii.The distribution of observations in a dot plot or a bar chart or histogram.

Related to the performance evaluation process in various fields, we have found several papers that were implemented AHP approach as an evaluation method. Kadarsah (2007) developed a framework of measuring key performance indicators for decision support in higher education institution. The framework is based on key success factors related to education institution sustainability which includes academic, research and supporting key performance indicators (KPI). AHP is used for weighted each KPI. As results, measurement values that reflect KPI scores are visualized in form of "wheel-shape". The developed framework contributes in measuring and explaining HLI success using multi dimensions of KPI.

In order to improve human resources management, Albayrak and Erensal (2004) applied AHP approach to solve the human performance improvement problem. This research presents a model which illustrates the relations and importance between human performance improvement and the style of management. The AHP was used for the purpose of structuring and clarifying the relations and important for both stated components. At the end of this research, they concluded the best management style in improving human performance is management by values.

Rafikul and Shuib (2005) also discussed about measurement of human performance system. An applied research type is discussed about the performance of individuals against organizational goals determines whether the organization meets its goals or otherwise. The criteria are measured based on quantity/quality of the work, planning, initiatives, teamwork, communication and several related external factors. Each of these criteria has been divided into three subcriteria and this developed framework was implemented to analyze real evaluation data. Overall ranking of employees has been obtained using the absolute measurement procedure of AHP.

Another research related to AHP application was innovated by Frair (1995). This research pursued an employment of AHP to assess the contributions of engineering student team members to the team effort. On the other hand, a description of AHP for teams within a production engineering class is also clearly illustrated. The minimal success of traditional student questionnaire to asses team performance is describe followed by a description of the what appears to be more meaningful results when AHP approach is used.

**3. General MCDM Design Model and Methodology for Teaching Performance Evaluation**

As stated in the comprehensive literature review in the previous section, it can conclude that MCDM is a powerful tool used widely for evaluating and ranking problems containing multiple, usually conflicting criteria (Pomerol and Barba Romero, 2000).

On the other hand, the MCDM approaches generally enable us to structure the problem clearly and systematically. With characteristic, decision makers have the possibility to easily examine the problem and scale it in accordance with their requirement.

During the literature survey, we have encountered few researches on the IHL instructors teaching performance evaluation process which is used any related MCDM method, although there are many studies related to the teaching performance evaluation procedure. Therefore, we propose a formal MCDM based approach, an AHP. The widest used tool for MCDM and also its features and ability to produce the most appropriate analyzed results (Liu et al. 2007) are the main factors of an AHP were chosen. On the other hand, this research focuses on implementation of AHP approach in HLI instructors teaching performance evaluation process which considering highly subjective of multiattribute criteria.

The teaching performance evaluation procedure of this particular study consists of three major steps as summarized in Fig. 3. They are includes

***Step 1:*** Identifying the teaching performance evaluation criteria that are considered the most important for evaluation purposes.
***Step 2:*** Constructing the evaluation criteria hierarchy and calculating weights by applying AHP approach.
***Step 3:*** Making a comparison between those two methods to achieve the final ranking results.

Regarding the suggested evaluation procedure above, the identification criteria and sub-criteria are most of the important part and must be clearly selected. As shown in Fig. 4, there are four determined criteria that must be considered during the teaching evaluation process for IHL instructors; preparation, organization, delivery and effectiveness.

Based on Fig.4 below, there are criteria levels which include four criteria. Beside that, there is also an alternative level which may consist of a group of HLI instructors which become as alternatives or candidates. This group may vary from certain departments, faculties/schools or as a whole of instructors groups in HLI.

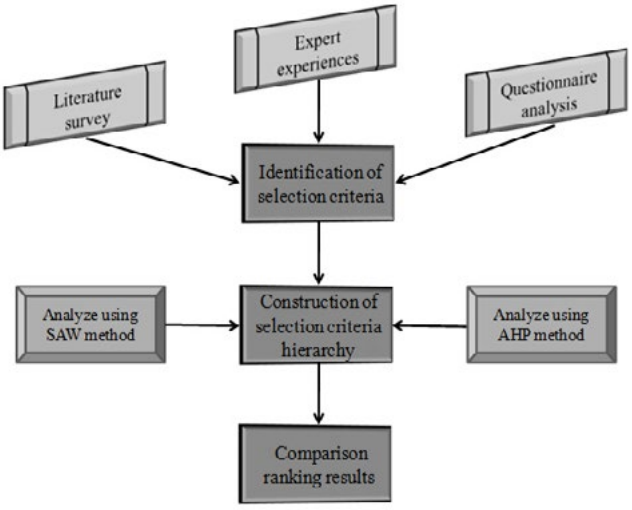


Fig. 3 General framework of evaluation procedure

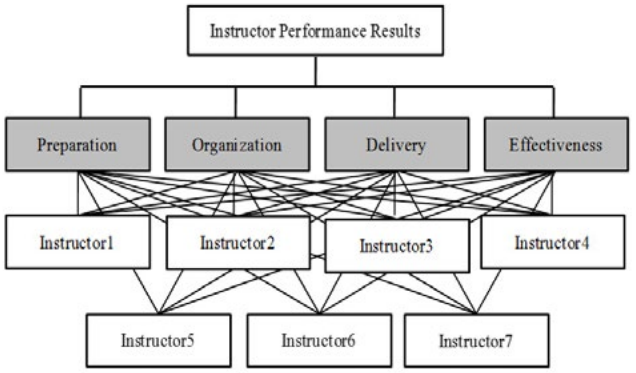


Fig. 4 Hierarchy of criteria for teaching evaluation procedure

**A. Teaching Performance Evaluation through MCDM Method**

Evaluating process for IHL instructors can be classified as MCDM problem. There are two kinds of methods for solving a MCDM problem includes compensatory and non-compensatory ones (Hwang and Yoon, 1981). The non-compensatory method does not permit tradeoffs among attributes. The MCDM techniques in this category are simple, but may not be suitable for the teaching performance evaluation. In contrast, the compensatory method allows tradeoffs among attributes. A slight decline in one attribute is accepted if it is compensated by some enhancement in one or more other attributes. These are three subgroups in this category; scoring, compromising and concordance methods.

We have selected the most appropriate approach from compensatory method as a guide line for developing suggested evaluation model, that is, an AHP approach. The next subsection is spent to discuss an implementation of selected method.

In order to illustrate on how to evaluate the teaching performance of IHL instructors and rank each instructor to the appropriate band, we use a sample situation below as a case study.

*There is one small department that consists of seven instructors which are selected to teach several university subjects for one semester. At the end of the particular semester, the university management decides to pursue the teaching performance evaluation of all instructors for teaching and learning improvement. University management considers four major criteria including preparation, organization, delivery and effectiveness. In addition, AHP approach was selected in this evaluation case.*

**B. The AHP Approach**

Based on the discussed system design in the previous section, AHP approach is used as an MCDM method to design teaching performance evaluation framework. In this section, the suggested teaching performance evaluation system will be presented general steps of AHP approach as a guide line.

Considering the general procedure that was discussed in the literature review section, we used the hierarchy of criteria for teaching evaluation procedure which was previously designed in Section 3.

After decomposing the problem into a hierarchy notation, alternatives at the given hierarchy are compared in pairs to assess their relative preference with regard to each criterion at the higher level. A scale is needed to represent the varying degrees of preference. Saaty (1980) established a scale in Table 4, where 9 is upper limit and 1 is lower limit and a unit difference between successive scale values is used. Adapted from Table 2, pair-wise comparison scale for AHP preference is summarized as shown below.

Table 4 Scores of Evaluations in Terms of Each Criterion

| Numerical Assessment | Linguistic Meaning |
|---|---|
| 1 | Equal preferred |
| 3 | Moderately preferred |
| 5 | Strongly preferred |
| 7 | Very strongly preferred |
| 9 | Extremely preferred |
| 2, 4, 6, 8 | Intermediate values of preferred |

$$S=\left\{\frac{1}{9},\frac{1}{8},\frac{1}{7},\frac{1}{6},\frac{1}{5},\frac{1}{4},\frac{1}{3},\frac{1}{2},1,2,3,4,5,6,7,8,9\right\}$$

A comparison is done among the elements on the same level of the goal hierarchy. In a comparing process, a value V from the scale is assigned to the comparison result of two criteria P and Q at first, and then the value of comparison of Q and P is reciprocal value of V. The value of the comparison of P and P is 1. Following these rules, a matrix of attributes is calculated through finding the eigenvector associated with the maximal eigenvalue of this matrix. A practically used algorithm is the following one.

i. Normalize each column by dividing each cell by the column total.
ii. Sum each of the rows into a new column.
iii. Normalize the new column by dividing each value by the sum of the column.
iv. Represent of the normalize column eigenvector, which contains the weight of each attribute.

The pair-wise comparison value of selected criteria is presented in Table 5 below.

Table 5 Pair-Wise Comparison Values of Criteria

| Criterion | Preparation | Organization | Delivery | Effectiveness |
|---|---|---|---|---|
| Preparation | 1 | 5 | 3 | 8 |
| Organization | 1/5 | 1 | 1/6 | 2 |
| Delivery | 1/3 | 6 | 1 | 6 |
| Effectiveness | 1/8 | 1/2 | 1/6 | 1 |

Table 6 shows a normalized pair-wise comparison value matrix of criteria and Table 7 shows the weight ranking for each criterion.

Table 6 Normalized Pair-Wise Comparison Values of Criteria and Their Weight

| Criterion | Preparation | Organization | Delivery | Effectiveness | Sum | Weight |
|---|---|---|---|---|---|---|
| Preparation | 0.518 | 0.579 | 0.574 | 0.541 | 2.212 | 0.553 |
| Organization | 0.091 | 0.072 | 0.089 | 0.081 | 0.334 | 0.084 |
| Delivery | 0.339 | 0.301 | 0.281 | 0.328 | 1.249 | 0.312 |
| Effectiveness | 0.052 | 0.048 | 0.056 | 0.049 | 0.205 | 0.051 |

The following step is to ranks the relative importance between each pair of alternatives in term of a criterion. Based on the calculated results, we notice that, a preparation criterion with 0.553 is the highest weight and followed by delivery with 0.312, organization with 0.084 and effectiveness as a less weigh which is 0.051 scores. The illustrated graph is shown in Fig. 6.

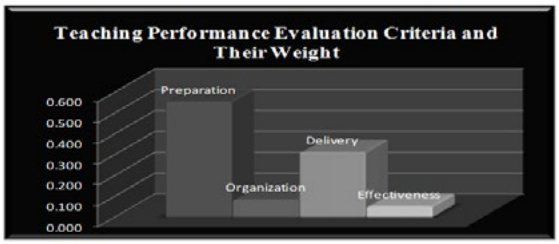


Fig. 6 Teaching performance evaluation criteria's and each weight

An index of consistency ratio (CR) can be used to measure consistency of an n-order square decision matrix. In AHP, the recommended threshold for CR is less than or equal to ten percent or 0.1, when the value of a CR is lower then this value; the decision matrix is accepted and can be applied to making decisions. Otherwise, the matrix is inconsistent and the pair-wise comparison weights must be revised.

The calculation of CR of a pair-wise comparison matrix is implemented by taking the following detail method innovated by Saaty (1980).

i. Lam max=the maximum eigenvalue (Perron root) of the matrix = Lmax =
ii. C.I. = Consistency Index = ( -n)/(n-1)
iii. R.I. = Random Index. For each matrix of size n, Saaty's team generated random matrices and computed their mean C.I. value and called it the Random Index.
iv. C.R. = Consistency Ratio = (C.I.)/(R.I). A value less than or equal to 0.1 is acceptable. Larger values require the decision maker to reduce the inconsistency by revising judgments.

Before this following method can be done, the computation of the n-th root of the product of the values in each row must be calculated first. This technique called the geometric mean based on the following, sample data, [a1, a2, ……, an],

$$GeoMean = \left( \prod_{i=1}^{n} a_i \right)^{1/n}, \quad (2)$$

Beside that, a priority vector for each row also must be calculated using the following formula:

$$priority\ vector = \frac{GeoMean}{\sum GeoMean} \quad (3)$$

Based on the pair-wise comparison value on Table 5, = 4.223 and the C.I. = 0.074. Consistency ratio = 0.083 which is < 0.1, therefore the judgments are consistent.

Similarly, we make a pair-wise comparison of the alternatives (instructors-INS) with respect to the four criteria in the higher level. The comparison and weights results are illustrated in Tables 7-10.

Table 7 Normalized Pair-Wise Comparison Values of Preparation Criteria and their Weight (*C.R. = 0.097*)

| Alternative | INS1 | INS2 | INS3 | INS4 | INS5 | INS6 | INS7 | Sum | Weight |
|---|---|---|---|---|---|---|---|---|---|
| INS1 | 0.116 | 0.164 | 0.156 | 0.160 | 0.126 | 0.151 | 0.143 | 1.016 | 0.145 |
| INS2 | 0.045 | 0.031 | 0.048 | 0.047 | 0.042 | 0.038 | 0.033 | 0.284 | 0.041 |
| INS3 | 0.322 | 0.303 | 0.285 | 0.289 | 0.340 | 0.303 | 0.337 | 2.179 | 0.311 |
| INS4 | 0.315 | 0.303 | 0.286 | 0.278 | 0.323 | 0.293 | 0.326 | 2.125 | 0.304 |
| INS5 | 0.119 | 0.136 | 0.144 | 0.146 | 0.101 | 0.147 | 0.104 | 0.897 | 0.128 |
| INS6 | 0.035 | 0.026 | 0.030 | 0.030 | 0.031 | 0.023 | 0.025 | 0.201 | 0.029 |
| INS7 | 0.048 | 0.037 | 0.050 | 0.049 | 0.039 | 0.045 | 0.030 | 0.298 | 0.043 |

Table 8 Normalized Pair-Wise Comparison Values of Organization Criteria and their Weight (*C.R. –0.098*)

| Alternative | INS1 | INS2 | INS3 | INS4 | INS5 | INS6 | INS7 | Sum | Weight |
|---|---|---|---|---|---|---|---|---|---|
| INS1 | 0.126 | 0.127 | 0.118 | 0.116 | 0.116 | 0.114 | 0.127 | 0.126 | 0.127 |
| INS2 | 0.288 | 0.257 | 0.282 | 0.265 | 0.279 | 0.291 | 0.273 | 0.288 | 0.257 |
| INS3 | 0.261 | 0.239 | 0.270 | 0.253 | 0.268 | 0.271 | 0.232 | 0.261 | 0.239 |
| INS4 | 0.167 | 0.197 | 0.191 | 0.204 | 0.200 | 0.188 | 0.201 | 0.167 | 0.197 |
| INS5 | 0.051 | 0.052 | 0.037 | 0.040 | 0.034 | 0.035 | 0.045 | 0.051 | 0.052 |
| INS6 | 0.077 | 0.098 | 0.074 | 0.092 | 0.076 | 0.072 | 0.080 | 0.077 | 0.098 |
| INS7 | 0.030 | 0.030 | 0.028 | 0.031 | 0.027 | 0.029 | 0.022 | 0.030 | 0.030 |

Table 9 Normalized Pair-Wise Comparison Values of Delivery Criteria and their Weight (*C.R. =0.090*)

| Alternative | INS1 | INS2 | INS3 | INS4 | INS5 | INS6 | INS7 | Sum | Weight |
|---|---|---|---|---|---|---|---|---|---|
| INS1 | 0.227 | 0.276 | 0.316 | 0.276 | 0.295 | 0.271 | 0.247 | 1.908 | 0.273 |
| INS2 | 0.207 | 0.172 | 0.201 | 0.212 | 0.195 | 0.222 | 0.207 | 1.415 | 0.202 |
| INS3 | 0.239 | 0.229 | 0.181 | 0.218 | 0.278 | 0.211 | 0.218 | 1.574 | 0.218 |
| INS4 | 0.132 | 0.123 | 0.127 | 0.101 | 0.120 | 0.137 | 0.146 | 0.885 | 0.126 |
| INS5 | 0.091 | 0.079 | 0.075 | 0.080 | 0.061 | 0.072 | 0.086 | 0.544 | 0.078 |
| INS6 | 0.070 | 0.082 | 0.063 | 0.073 | 0.059 | 0.052 | 0.063 | 0.461 | 0.066 |
| INS7 | 0.035 | 0.041 | 0.037 | 0.040 | 0.041 | 0.036 | 0.033 | 0.263 | 0.038 |

Table 10 Normalized Pair-Wise Comparison Values of Effectiveness Criteria and their Weight (*C.R. –0.076*)

| Alternative | INS1 | INS2 | INS3 | INS4 | INS5 | INS6 | INS7 | Sum | Weight |
|---|---|---|---|---|---|---|---|---|---|
| INS1 | 0.040 | 0.041 | 0.044 | 0.042 | 0.046 | 0.046 | 0.043 | 0.302 | 0.043 |
| INS2 | 0.056 | 0.047 | 0.058 | 0.063 | 0.063 | 0.059 | 0.058 | 0.405 | 0.058 |
| INS3 | 0.111 | 0.097 | 0.077 | 0.088 | 0.111 | 0.109 | 0.108 | 0.702 | 0.100 |
| INS4 | 0.133 | 0.140 | 0.113 | 0.108 | 0.143 | 0.138 | 0.132 | 0.906 | 0.129 |
| INS5 | 0.193 | 0.231 | 0.198 | 0.185 | 0.139 | 0.207 | 0.199 | 1.353 | 0.193 |
| INS6 | 0.199 | 0.180 | 0.234 | 0.239 | 0.201 | 0.167 | 0.190 | 1.409 | 0.201 |
| INS7 | 0.267 | 0.264 | 0.276 | 0.275 | 0.297 | 0.274 | 0.270 | 1.924 | 0.275 |

**C. Results of the AHP Application**

In order to determine the best performance of instructors, Saaty's (1980) AHP was applied to determine performance point values for this applied research as in Fig. 4. The priority rankings for each instructor were determined from a hierarchy based on four criteria and seven alternatives. The selected criteria were compared on a pair-wise basis. This produced a ranked score for each instructor on each of the criteria and ranked the score for each criterion. Multiplying these scores provides a summary score for each the criteria are determined, each of the alternative was compared on the basis of the criteria. The eigenvalue of the square matrix, the preference vector, is then computed to determine the relative ranking of the four criteria in selecting the best evaluation point.

Table 11 Relative Scores for Each Objective and Original Weight of the Criteria

| Criteria/Alternative | INS1 | INS2 | INS3 | INS4 | INS5 | INS6 | INS7 | Weight |
|---|---|---|---|---|---|---|---|---|
| Preparation | 0.145 | 0.041 | 0.311 | 0.304 | 0.128 | 0.029 | 0.043 | 0.553 |
| Organization | 0.121 | 0.276 | 0.259 | 0.192 | 0.042 | 0.081 | 0.028 | 0.084 |
| Delivery | 0.273 | 0.202 | 0.218 | 0.126 | 0.078 | 0.066 | 0.038 | 0.312 |
| Effectiveness | 0.043 | 0.058 | 0.100 | 0.129 | 0.193 | 0.201 | 0.275 | 0.051 |
| **Overall Scores** | **0.178** | **0.112** | **0.267** | **0.230** | **0.109** | **0.054** | **0.052** | |

Based on the preference vectors given in Table 11 above, the four evaluation criteria are ranked, thus, preparation is the most important criterion for teaching performance evaluation whereas effectiveness is of least consider. The final results which show the evaluation ranking are illustrated in the graph below which shows the best performance in HLI instructors teaching performance activities for this case is Instructor3 which 0.267 scores and Instructor7 with 0.052 which is the lowest point of teaching performance evaluation results.

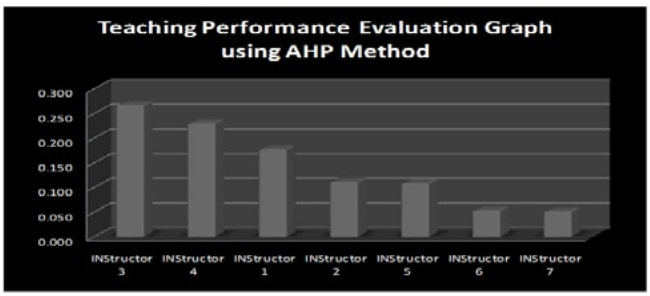


Fig. 7 Teaching performance evaluation results and each weight

With an implementation of the AHP approach to the sample case, which is considering four evaluation criteria includes; preparation, organization, delivery and effectiveness, we find that the produced results are consistently calculated. It is proof by a consistency ratio value which is less then 0.1. On the other hand, the listed teaching performance evaluation criteria's of HLI instructors are highly in subjective evaluation term. Based on the constructed sample of case study, we find that the discussed ability and features of AHP are truly presented here. With two levels of hierarchical process; criteria's level and alternative level, the main objective of this research is achieved which preference ranking of alternatives (in this case, instructor's) is clearly concluded and highly meaningful.

**4. Conclusion**

This paper introduces a model, based on Saaty (1980) called AHP which determines the priority weight for teaching performance evaluation and has examined the critical factors that effect HLI instructor's performance. The AHP methodology could help assess relevant criteria critically and

logically and to assist in making sensible decisions. In addition, the proposed AHP framework that might assist in providing direction for evaluating instructors teaching performance point which almost becomes as needs in HLI.

Based on the eigenvalues computed by the proposed framework, we can conclude that preparation criterion is the most important in order to determines the IHL instructors teaching performance evaluation. The preparation includes several activities such as discussing teaching activities plan and clear assessment, and exhibited knowledge of the subject. However, effectiveness criterion was the minimum consideration for teaching performance evaluation process. In fact, the relative ranking is preparation, organization, delivery and effectiveness.

Based on the discussed case study which experimentally used AHP for evaluating HLI instructors was teaching performance, we found that Instructor3 shows the best instructor performance on the basis of the four discussed criteria's. On the other hand, the minimum performance point for this case study is Instructor7 which may to improve his/her related teaching activities.

The findings of this study point to the need for HLI instructors itself and also their management to understand the impacts that have to consider for improve the overall institution performance. Through the human performance analysis process, especially related to the teaching and learning environment, managers or persons who are in the top management will be able to make better decisions about the work organization and management strategies in order to enhance their education business. As strategy tool, the understanding of human performance will help to identify the right management technique for managing valuable human resources group that fit into intuitional vision. Effective AHP method supported teaching performance evaluation process can play a significant role in helping intuitions retain and to improve their best critical talent.

This study is done to consider only internal criteria related to the teaching performance evaluation process. Therefore, also external criteria's can be considered for enhancing this research such as size of student and class, instructor's gender, class time slot and so on. Furthermore, artificial intelligence techniques can be used in this kind of applications. For instance, expert system may be used to enhance the explanation facilities or neural network techniques to evaluate and forecast the potential of instructor's teaching performances relating to different cases. Hence, it is important to note that the aim of the proposed evaluation system is not to replace the current system of evaluating performance but it may be used to strengthen and improve the present system of evaluation by providing additional information for appraiser to make decision in teaching performance evaluation and human resources management in an organization especially HLI.

**References**

[1] C. Kirkwood, Strategic Decision Making: Multiobjective Decision Analysis with Spreadsheets, Duxbury Press, Belmont, CA, 1997.

[2] C. L. Hwang and K. Yoon, Multiple Attribute Decision Making, Lecture Note in Economics and Mathematical System, Vol. 186. Berlin: Springer-Verlag, New York, 1981.

[3] E. Albayrak and Y. C. Erensal, Using Analytic Hierarchy Process (AHP) to Improve Human Performance: An Application of Multiple Criteria Decision Making Problem, Journal of Intelligent Manufacturing, Vol. 15, No. 4, pp 491-503, 2004.

[4] F. Colace, M. De Santo and A. Pietrosanto, Evaluation Models for E-Learning Platform: An AHP Approach, Proceedings, the 36th ASEE/IEEE Frontiers in Education Conference, San Diego, CA, pp 1-6, 2006.

[5] G. L. Geis, Evaluation: Definition, Problems and Strategies in Chris Knapper et al. Eds., Teaching is Important (Toronto: Clarke Irwin in Association with CAUT), 1997.

[6] J. C. Pomerol and R. S. Barba, Multicriteria Decision in Management: Principles and Practice, 1st Edition, Kluwer Academic Publishers, Norwell, 2000.

[7] J. Malczewski, GIS and Multicriteria Decision Analysis, John Wiley and Sons, New York, NY, 1999.

[8] J. Malczewski, Propagation of Errors in Multicriteria Location Analysis: A Case Study. Multiple Criteria Decision Making, Proceedings, the Twelfth International Conference. Hagen (Germany), G. Fandel, and T. Gal, Springer-Verlag, pp 154-165, 1997.

[9] J. Perez, Some Comments on Saaty's AHP, Management Science, Vol. 41 No. 6, pp 1091-1095, 1995.

[10] L. Frair, Student Peer Evaluations using the Analytic Hierarchy Process Method, Proceedings, the 1995 Frontiers in Education Conference, Atlanta, GA, pp 1-5, 1995.

[11] Rafikul Islam and Shuib bin Mohd Rasad, Employee Performance Evaluation by AHP: A Case Study, Proceedings, the dISAHP 2005, Honolulu, Hawaii, USA, pp 163-176, 2005.

[12] R. M. Shen, Y. Y. Tang and T. Z. Zhang, The Intelligent Assessment System in Web Based Distance Learning Education, Proceeding, 31st ASEE/ IEEE Frontiers in Education Conference, Vol. 1, pp 7-11, 2001.

[13] R. T. C. Boyd, Improving Teacher Evaluations Practical Assessment, Research & Evaluation, Electronic Journal, American Institutes for Research, available at: http://ericae.net/pare/getvn.asp?v=1&n=7 (accessed 15 June 2008), Vol. 1, No 7, 1989.

[14] S. Kadarsah, Framework of Measuring Key Performance Indicators for Decision Support in Higher Education Institution, Journal of Applied Science Research, Vol. 3, No. 12, pp 1689-1695, 2007.

[15] S. K. Lai and L. D. Hopkins, The Meaning of Trade-off in Multiattribute Evaluation Methods; A Comparison Environment Planning B, Planning and Design, Vol. 16, No. 2, pp 155-170, 1989.

[16] T. L. Saaty, Mathematical Methods of Operations Research, Dover, New York, 1988.

[17] T.L. Saaty, The Analytical Hierarchy Process: Planning, Priority Setting, Resource Allocation, New York: McGraw-Hill, 1980.

[18] Y. W. Liu, Y. J. Kwon and B. D. Kang, A Fuzzy AHP Approach to Evaluating E-Commerce Websites, Proceedings of the Fifth International Conference on Software Engineering Research, Management and Applications, Busan, Korea, pp 114-122, 2007.

[19] Y. Zhu and A. P. Buchman, Evaluating and Selecting Web Sources As External Information Resources of a Data Warehouse, Proceedings, the Third International Conference on Web Information Systems Engineering, pp 149, 2002.

[20] W. J. McKeachie, Student Ratings: The Validity of Use, American Psychologist, Vol. 52, No. 11, pp 1218-1225, 1997.

Cite this article as: Azizul Azhar Ramli, Shahreen Kasim, Mohd. Farhan Md. Fuzzee, Hairulnizam Mahdin, Teaching Performance Evaluation Framework: An Analytic Hierarchy Process Approach/ Acta Inform. Malay. (1) (2017) 01-06