The Improvement Technique in Assessing Teaching Skills using Fuzzy Weighted Grading Method

S Andayani
Mathematics Education Department, Yogyakarta State University, Karangmalang Yogyakarta, Indonesia, 55281
andayani@uny.ac.id

Abstract. The assessment of education is the process of providing the decision on the performance and quality of student achievement. Many cases indicate that the educational assessment involves many criteria of uncertain value, and therefore fuzzy assessment studies have grown rapidly in recent years. In this paper, a method is proposed to improve the assessment of the basic teaching skills of prospective teachers in a Micro-teaching course by using a fuzzy grading system. This method is also based on group decision-making (GDM) because it involves several decision makers who give preferences in the assessment process and negotiate to determine the criteria and sub-criteria for the assessment. The result of the method is a fuzzy value which is represented by a letter grade for the Micro-teaching course and letter grade for each of the basic teaching skills. By using this method, students may have the same final grade, but their mastering of teaching skills is very different.

1. Introduction
The assessment of education is related to the process of decision-making on the quality or achievement of student performance [1]. In the course handled by more than one lecturer, the assessment process is a group decision-making (GDM).

In GDM, the group will be involved in the decision-making process such as the preparation of an alternative list of criteria that is worth considering and selecting the criteria to be used [2]. The decision-making process in the assessment is in line with the activities and process characteristics of GDM, which are: (1) the decision-making process is a joint activity involving a group of people who have the same or similar status; (2) the output of the negotiations depends in part on the knowledge, opinion, and decisions of the participants; (3) the output also depends on the group composition and decision-making process used by the group; (4) disagreements are organized by way of member rankings, negotiations or arbitration [2].

The decision-making process in the assessment often involves criteria whose value cannot be ascertained. Therefore, many studies on educational assessments use fuzzy set theory. Several studies have shown that fuzzy-based evaluation systems are more reliable [3][4][5][6]. The proposed system can also reward students' performance fairly [7]. The ability of students in the competence of a course is assigned a specific value, so the superiority of each student can be known in different fields.

One of the group decision-making models based on fuzzy set theory is the method proposed by Ma and Zhou [8]. In the method, an integrated fuzzy set of approaches is presented to assess student-centered learning outcomes. This approach uses 4 steps: (1) determination of a set of assessment criteria...
conducted by lecturers together with students; (2) select criteria from a set of proposed criteria, using the Analytic Hierarchy Process method; (3) determine the weight of the selected criteria; (4) evaluate student learning outcomes using selected criteria. The GDM process in this method is done in steps 1 and 2. Based on the selected collective assessment criteria, student learning outcomes are evaluated based on the fuzzy grading system. But, in the method there is no step to determine the value of student performance in the criteria that are components of the assessment. Thus, the results of the assessment cannot describe the performance of students in each criterion that determine the final outcome of the assessment.

An evaluation method that explicitly illustrates the student's ability in the criteria determining the final outcome of the assessment is proposed by Mossin et al. [7]. The method proposes the use of fuzzy set techniques in the evaluation process within the learning areas of industrial automation systems. The evaluation process is based on three criteria and the result is the classification of students' ability in learning and industrial automation system.

One of the compulsory subjects in higher education involving a group of lecturers is Micro-teaching. This course prepares prospective teachers to have basic teaching skills. The basic competencies of teaching taught in Micro-teaching consist of the preparation of lesson plans, limited basic teaching competencies, and integrated basic teaching competencies. In this course, students practice teaching in small groups/classes consisting of 8-10 students. When a student is in charge of teaching, the other students act as students (peer teaching). In one semester, prospective student teachers practice teaching as much as 3-5 times. Assessment of the teaching practice was conducted by more than one lecturer using observation sheets of various types of basic teaching skills.

Although it involves several basic teaching skills as assessment criteria, the Micro-teaching assessment does not take into account the weight of the importance of each criterion. In this case, each criterion is considered to have equal weight in the final value calculation. It can be regarded as an unfair assessment, because each student may have different strengths in different skills but get the same final score. Therefore, an improvement technique to assess the teaching skills of prospective teacher is needed to accommodate the requirement.

It has been suggested that teaching skills cannot be assessed in a satisfactory way [9]. In principle, teaching skills shall be assessed as well as academic skills. A documentation of the qualifications is required and the assessment shall be based on clearly defined criteria. In order to make valid and reliable assessments the scope and meaning of teaching skills needs to be clearly formulated. Instead of the final grade of a course in teaching skills, the achievement of each type of skill should be explored to evaluate the strengths and weaknesses of prospective teachers.

Although the proposed model is aimed at assessing teaching skills, the model can be used more generally for the assessment of other courses using the same principles. The principles are that assessment is done by many teachers, against many criteria and each criterion has a weight. Thus, the model can also be used to assess the character of the student, by adding a character assessment as one of the assessment criteria.

Based on the prior methods, the improvement technique to assess the basic teaching skills is proposed as follows.

2. The proposed improvement technique
The assessment / assessment of basic teaching skills of prospective teachers in Micro-teaching involves decision-makers (D₁, D₂, ..., Dₙ) in accordance with the type of basic skills to be assessed (K₁, K₂, ..., Kₖ). Each type of teaching has different sub-criteria. Decision makers have a right to give different weight to each sub-criteria in their respective authority. The assessment process of the m student (A₁, A₂, ..., Aₘ) is done b times in one semester. At the time of the assessment, according to the decision maker scores each sub-criteria against the student who practices the teaching.

There are 3 main steps used in the proposed technique to determine the final grades of prospective teachers in Micro-teaching. The three steps are as follows.

1. Aggregate decision makers' preferences to score student performance of basic teaching skills
2. Determine the weight of each kind of teaching skill by using AHP (Analytic Hierarchy Process) method, based on the result of negotiation from the decision maker.
3. Determine the final value and value of each type of teaching skill by using fuzzy grading system.

2.1. Aggregate preferences to determine performance values
The preferences of decision makers are represented in the decision matrix, and then aggregated to get the value of the performance of each type of teaching skill. The average performance value of each $i^{th}$ alternative in the $t^{th}$ type of teaching skill of $b$ assessment can be determined by

$$u_i^t(x) = \frac{1}{b} \sum_{s=1}^{b} u_i^{ts}(x)$$

where $u_i^{ts}$ is score of the $i^{th}$ skill performance in the $s^{th}$ assessment; $s = 1, 2, ..., b$. $w_j^t$ is the weight of the $j$ sub-criteria of the type of skill $t$. $r$ is the normalization of the decision matrix:

$$r_{ij}^t(x) = \frac{x_{ij}}{\max_x x_j}$$

The above steps are adopted from Simple Additive Weighting (SAW) method, because SAW is considered simple and widely used in solving the problem of multiple attribute decision making [10].

2.2. Determine the weight of each criterion
There are many methods for calculating the weight of criteria for decision making, including Analytic Hierarchy Process (AHP) and Analytic Network Process (ANP). Both of these methods use pair-wise comparison matrix (PCM) as the initial step of weight calculation. In cases where the criteria involved are quite large, PCM is quite complicated, so there is a need for help to set up the PCM to remain consistent [11].

In the proposed method, the weight of the assessment criteria is determined by using AHP. In AHP, the degree of importance among criteria can be compared on a scale of 1-9 to a pair-wise comparison matrix $A$ using a scale proposed by Saaty [12].

2.3. Determine the final value using the fuzzy grading system
Generally, the final grade of a course is presented on a letter scale, for example $G=[g_1, g_2, ..., g_h]$ where $g_h$ is fuzzy grade in letter and $h$ is the number of letter scale. For example: $g_1= A$: Excellent; $g_2 = B$: Good; $g_3= C$: Fair; $g_4= D$: Poor.

Membership functions must be defined to transform the numeric values into letter values. The performance scores in the scale number 100, $u_i$, have membership degrees $\mu_{g_h}(u_i)$ in accordance with the grade $g_h$. A fuzzy mapping function can be defined, which combines the membership function together to obtain a letter value.

$$f : A \rightarrow \Gamma(G)$$

$$A_i \rightarrow (\mu_{g_1}(u_i), \mu_{g_2}(u_i), ..., \mu_{g_h}(u_i))$$

where $u_i$ is the performance score of student $A_i$

Performance score in the scale of the letters of the $i^{th}$ student on the type of skill $t$ is determined from the best degree of membership of the skill,
The writing of $\mu_{gh}(u'_i)$ in equation 4 can be simplified by using a notation $e_{ij}$ a fuzzy evaluation matrix $E$, expressed as follows.

$$E = \begin{bmatrix}
e_{11} & e_{12} & \cdots & e_{1d} \\
e_{21} & e_{22} & \cdots & e_{2d} \\
\vdots & \vdots & \ddots & \vdots \\
e_{n1} & e_{n2} & \cdots & e_{nd}
\end{bmatrix} = \begin{bmatrix} e_{ij} \end{bmatrix}$$

(5)

where $i=1,2,\ldots,n$ is the number of skill type

and $j=1,2,\ldots,d$ is the number of fuzzy grade / categories in letters.

The final value in the letter scale can be determined by using the Y vector

$$Y = \omega \odot E = (\omega_1, \omega_2, \ldots, \omega_n) \odot \begin{bmatrix} e_{ij} \end{bmatrix}$$

(6)

where $y_i = (\omega_1 \cdot e_{i1}) \oplus (\omega_2 \cdot e_{i2}) \oplus \cdots \oplus (\omega_n \cdot e_{ni})$

(7)

where $\omega$ is the matrix of weight of assessment criteria.

Operator $\cdot$ is an algebraic product, $c = a \cdot b \rightarrow c = ab$; and operator $\oplus$ is bounded sum, $c = a \oplus b \rightarrow c = \min(1, a+b)$

3. Result and Discussion

In the application of the proposed method, the assessment of basic teaching skills in Micro teaching is conducted by 5 decision makers in the different field of expertise. Assessment criteria are summarized from integrated and limited basic teaching skills, so that 5 teaching skill types with 27 assessment sub-criteria are shown in Table 1. In giving the assessment, the lecturer gives score in the form of score 1-4 on the sub-criteria according to his expertise, based on the observation when the students do the teaching practice using the observation sheet.

3.1. Aggregate preferences and calculate performance values

Table 2 shows the average performance score of students in each type of teaching skill which is the result of aggregation calculation of decision makers' preferences from 4 times assessments. These values are calculated by using equations (1) and (2). Value is presented on a scale of 100.
15. Skill of conducting variations
16. Skills to guide discussion
17. Classroom management skills
18. Match the media with learning materials
19. Streamline student learning
20. Skill to prepare/create media
21. Skill to use media
22. Creativity chooses materials and compiles the media

Expert of Media
Mastering the media (MED)

Expert of evaluation
Mastering the evaluation (EVA)

Table 2. The average score of 4 times performance

| Student# | MAT  | LP  | PDG  | MED  | EVA  |
|----------|------|-----|------|------|------|
| 1        | 78.96| 80.10| 73.23| 80.21| 80.42|
| 2        | 78.85| 80.94| 80.63| 86.67| 84.17|
| 3        | 71.25| 79.48| 76.25| 74.79| 83.75|
| 4        | 93.75| 91.04| 92.19| 88.75| 87.50|
| 5        | 68.96| 77.71| 73.33| 79.58| 77.92|
| 6        | 54.27| 75.42| 64.27| 67.92| 67.92|

3.2. Determine the weight of each type of teaching skill
The important weight of each type of teaching skill is obtained by forming a pair-wise comparison matrix A, derived based on the results of negotiations between the lecturers involved in the assessment. Table 3 presents comparisons of (n-1) the importance level between types of teaching skills.

| Mastering the material | The meaning |
|------------------------|-------------|
| 1                      | equally important |
| 1/3                    | Mastering the material is weakly more important than the preparation of LP |
| 1/5                    | Mastering the material is more important than mastery of pedagogy |
| 1/7                    | Mastering the material is strongly important than the mastery of the media |
| 1/9                    | Mastering the material is absolutely more important than media mastery |

Therefore, a pair-wise comparison matrix A is obtained as follows.

\[ A = \begin{bmatrix} 1 & 3 & 5 & 9 & 7 \\ 1/3 & 1 & 5/3 & 3 & 7/3 \\ 1/5 & 3/5 & 1 & 9/5 & 7/5 \\ 1/9 & 1/3 & 5/9 & 1 & 7/9 \\ 1/7 & 3/7 & 5/7 & 9/7 & 1 \end{bmatrix} \]
The priority weight of \( \omega_i \) \( i=1,2,\ldots,5 \) obtained from pair-wise comparison matrix \( A \) is
\[
0.5595 \quad 0.1865 \\
0.1119 \quad 0.0622 \quad 0.0799
\]

3.3. Determine the final value and value of each type of teaching skill

The output of this method is the final grade of the Micro teaching and the value of student mastery on each type of basic teaching skill. The letter value uses nine (9) categories which are determined based on the triangular membership function. The representation of the function is shown in the graph in Figure 1.

The membership degree of performance value of each student in each type of skill is calculated, and the results are stored as the matrix \( E \). Five (5) rows of matrix \( E \) indicate the types of teaching skills, respectively are MAT, LP, PDG, MED and EVA. Nine (9) columns of matrix \( E \) from left to right represent values of the letters \( E \), \( D \), \( C \), \( C^+ \), \( B^- \), \( B \), \( B^+ \), \( A^- \) and \( A \) respectively.

The final grade of the type of teaching skill is determined from the category with the highest degree of membership (equation 4). For example, in the matrix \( E_1 \) represents the performance value of student #1. From the matrix \( E_1 \) is known, the highest membership degree of mastering material (MAT) is in the 7th column (see row 1), which corresponds to a \( B^+ \) value. Thus, the value of mastering material for student #1 is \( B^+ \).

\[
\begin{align*}
E_1 &= \begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 1.0000 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0.9479 & 0.5521 & 0 \\
0 & 0 & 0 & 0 & 0 & 1.00 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0.8958 & 0.6042 & 0 \\
0 & 0 & 0 & 0 & 0 & 0.7917 & 0.7083 & 0
\end{bmatrix}
\end{align*}
\]

The final grade of micro teaching for each student is obtained using Equation 6 and 7. The end value of the student #1 is calculated in the following way.

\[
\begin{align*}
Y &= \omega \circ E \\
y_i &= (\omega_1 \cdot e_{1i}) \oplus (\omega_2 \cdot e_{2i}) \oplus (\omega_3 \cdot e_{3i}) \oplus (\omega_4 \cdot e_{4i}) \oplus (\omega_5 \cdot e_{5i}) \quad i=1,2,\ldots,9 \\
y_6 &= (\omega_1 \cdot e_{16}) \oplus (\omega_2 \cdot e_{26}) \oplus (\omega_3 \cdot e_{36}) \oplus (\omega_4 \cdot e_{46}) \oplus (\omega_5 \cdot e_{56}) \\
y_6 &= (0.5595 \times 0) \oplus (0.1865 \times 0) \oplus (0.1119 \times 1) \oplus (0.0622 \times 0) \oplus (0.0799 \times 0) \\
y_6 &= \min(1, 0+0+0.1119+0+0) \\
y_6 &= 0.1119
\end{align*}
\]

The final value of each type of teaching skill uses nine (9) categories which are determined based on the triangular membership function. The representation of the function is shown in the graph in Figure 1.

Figure 1. Membership function of the nine letter value
In the same way, $y_i; i=1,2,..,9$ obtained for the student \#1 as follows.

\[
Y = [0 \ 0 \ 0 \ 0 \ 0.1119 \ 0.8553 \ 0.1971 \ 0 \ 0]
\]

Based on these results, the maximum score of the matrix $Y$ is 0.8553, is in the 7th column, indicating the value $B^+$, so the final value of the student\#1 is $B^+$.

Let see the two matrices $E_1$ and $E_2$ as the final score of two students. Both of the students get the same final score, but have different levels of mastery on some types of teaching skills.

The matrix $E_1$ shows the membership degree of the value of each type of teaching skill of student\#1 in grade E (from left) to grade A (rightmost). The meaning of the matrix is:

- row-1 is the criterion of mastering the material in the $B^+$ category
- row-2 is the criterion for preparing LP in the $B^+$ category
- row-3 is the criterion of mastery of pedagogy in category B
- row-4 is the criterion of mastering the media in the $B^+$ category
- row-5 is the criterion of mastering the evaluation in the $B^+$ category

The matrix $E_2$ for the value of student\#2 is as follows.

\[
E_2 = \begin{bmatrix}
0 & 0 & 0 & 0 & 0.6875 & 0.8125 & 0 \\
0 & 0 & 0 & 0 & 0.1667 & 1.000 & 0 \\
0 & 0 & 0 & 0 & 0 & 0.1000 & 0.0833
\end{bmatrix}
\]

The meaning of the matrix $E_2$ is:

- row-1 is the criterion of mastering the material in the $B^+$ category
- row-2 is the criterion for preparing LP in the A- category
- row-3 is the criterion of mastery of pedagogy in category A-
- row-4 is the criterion of mastering the media in the A category
- row-5 is the criterion of mastering the evaluation in the A- category

The final value, $Y$, for student\#2 is

\[
0 \ 0 \ 0 \ 0 \ 0.7355 \ 0.3619 \ 0.0688
\]

Matrix $Y$ has the maximum score 0.7355 in column number 7, it means, the final grade of the student \#2 is $B^+$.

An example of the final score of the two students shows that both students get the same final score, $B^+$, but differ on the level of mastery on some types of skills. The second student has a better level of mastery on the preparation of LP, pedagogical mastery, media and evaluation. If it is noticed in the final value of processing scores, the value of the second student tends to mark A, which is indicated by the score on the mark A- is 0.3619 and on the mark A is 0.0688. While the first student score revolves around the mark B with a score of 0.1119 and at mark A- of 0.1971. The greater weight on the mastery of the material allows both students to get the same score, $B^+$, although the second student's mastery of the four basic types of teaching skills is better than the first student. Score of final value (matrix $Y$) which presents scores on all categories of letter values used will greatly assist decision makers in determining final mark of the competencies of students.

The result of the final score (matrix $Y$) shows the membership degree of the grade on all letter grades. The final value recommendation is determined on the basis of the greatest score. The final value processing matrix can be used as a consideration in providing a more comprehensive judgment, as it presents scores on all categories of letter values used. In the given example, the 1st and 2nd students are equally recommended to get the final score $B^+$, but the distribution of the 1st student's grades is at B, $B^+$ and A-, while the 2nd student is on the value $B^+$, $A^-$ and A.
The assessment criteria in this application are still limited to basic teaching skill, but other criteria can be developed which include the character of prospective teachers as assessment criteria. The character assessment will certainly involve a lot of uncertainty, so the role of fuzzy set theory will greatly help solve this problem [13].

4. Conclusion

The proposed method of assessment incorporates the principle of aggregating decision-making preferences from the SAW method, calculating the weights of basic teaching skills with AHP and treating the final value with the fuzzy grading system. In addition, the importance weight of each teaching skill gives added value to the final result of the assessment. The values in the E matrix indicating the level of student mastery in all types of basic teaching skills will also be helpful as a feedback material for improving the quality of basic skills of teaching students.

References

[1] Tay, M. K., Lim, P. C., and Jee, T. L. 2010. Enhancing fuzzy inference system based criterion-referenced assessment with an application. Proceedings 24th European Conference on Modelling and Simulation, ECMS.
[2] Turban E and Aronson J 2001 Decision support systems and intelligent systems 6th ed Prentice Hall New Jersey
[3] Hameed I A, and Sorensen CG 2010 Fuzzy Systems in Education: A More Reliable System for Student Evaluation Fuzzy Systems InTech accessed on 21 Maret 2011
[4] Capuano N, Loia N, and Orciuoli F 2016 A Fuzzy Group Decision Making Model for Ordinal Peer Assessment, IEEE Transactions on Learning Technologies.
[5] Jevšček M 2016 Competencies assessment using fuzzy logic Journal of Universal Excellence, 5(2) 187–202.
[6] Xiaojun Z, and Yunfeng W 2016 Quality Evaluation of Entrepreneur Education on Graduate Students Based on AHP-fuzzy Comprehensive Evaluation Approach, International Journal of Engineering Research & Science (IJOER) 2(1)
[7] Mossin EA, Pantoni R P, and Brandão D 2010 Students’ Evaluation based on Fuzzy Sets Theory. Fuzzy Systems, InTech accessed on 21 Maret 2011.
[8] Ma J, and Zhou D 2000 Fuzzy Set Approach to the Assessment of Student-Centered Learning. IEEE Transactions On Education 43(2)
[9] UPI Uppsala Universitet, Assessing Teaching Skills in Higher Education, Office for Development of Teaching and Interactive Learning, (UPI). Uppsala University Box 513 751 20 Uppsala
[10] Tseng G H and Huang J J 2011 Multiple attribute decision making, methods and applications. CRC Press. Boca Raton, FL.
[11] Andayani S, Hartati S, Wardoyo R, and Mardapi D 2016 Combining rating scale and FANP to determine the important weight of learning competency for students’ assessment, Proceedings of The 1st International Conference on Science and Technology 2015 (ICST-2015), 11–13
[12] Li Bo, Xuning P, and Bingquan B 2009. Modelling of Network Education Effectiveness Evaluation in Fuzzy Analytic Hierarchy Process International Conference on Networking and Digital Society IEEE.
[13] Ismail M, and Syaiful L 2015 Affective Assessment in Learning using Fuzzy Logic, Proceeding of IEEE Conference on e-Learning, e-Management and e-Services.