A new fuzzy peer assessment methodology for cooperative learning of students

Kok Chin Chai [a,*], Kai Meng Tay [a], Chee Peng Lim [b]

[a] *Faculty of Engineering, Universiti Malaysia Sarawak, Malaysia*
[b] *Centre for Intelligent Systems Research, Deakin University, Australia*

ARTICLE INFO

ABSTRACT

In this paper, a new fuzzy peer assessment methodology that considers vagueness and imprecision of words used throughout the evaluation process in a cooperative learning environment is proposed. Instead of numerals, words are used in the evaluation process, in order to provide greater flexibility. The proposed methodology is a synthesis of perceptual computing (Per-C) and a fuzzy ranking algorithm. Per-C is adopted because it allows uncertainties of words to be considered in the evaluation process. Meanwhile, the fuzzy ranking algorithm is deployed to obtain appropriate performance indices that reflect a student's contribution in a group, and subsequently rank the student accordingly. A case study to demonstrate the effectiveness of the proposed methodology is described. Implications of the results are analyzed and discussed. The outcomes clearly demonstrate that the proposed fuzzy peer assessment methodology can be deployed as an effective evaluation tool for cooperative learning of students.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Cooperative learning is an educational approach or strategy in which students work in small groups to help each other to learn academic content [1]. The importance of cooperative learning in engineering disciplines has been explained and highlighted in [2–4]. Cooperative learning plays an important role to improve students' soft skills, e.g., communication and teamwork, which is the essence in engineering studies [2]. While cooperative learning offers remarkable benefits (e.g., improving collaborative and critical thinking skills) to students at the tertiary level (i.e., the third stage of learning after graduating from the secondary school) [1], it is yet to be widely adopted owing to a number of practical challenges [5]. A search in the literature reveals that the assessment of an individual student in a group is not an easy task, since a group mark is often not a clear and fair reflection of each individual's effort [5–7]. Besides that, it is difficult for an instructor to closely monitor each student's efforts in a group; therefore it is not suitable for the instructor to assess each student's contribution [8,9]. To tackle these challenges, peer assessment has been introduced to evaluate each student's contribution in a group work [5–13]. Several successful case studies in peer assessment have been reported, e.g., in

civil engineering [5], biological sciences [10], primary mathematics education [11], and computer studies [13]. In addition, substantial evidence to show that peer assessment can lead to improvements in quality and effectiveness learning is available [14].

Generally, there are two types of peer assessment [10]: (i) involving students in a class to assess other students' work; (ii) involving students to assess the contribution/performance of other students within the same group. These two types of peer assessment can be further classified into two; i.e., formative or summative assessment [15,16]. The goals of formative assessment are to monitor students' learning capabilities, gather their ongoing feedbacks, and improve their learning experience [15,16]. On the other hand, summative assessment evaluates students' learning capabilities at the end of an instructional unit [15,16]. Typically, an instructor needs to decide whether to use the formative or summative form of peer assessment [14]. This paper focuses on summative assessment, which focuses on the outcome of a learning process [9]. The procedure for summative assessment is further detailed in Section 2.4.

Traditionally, the Likert scale (a numerical grading scale) is used in a way equivalent to psychological measurement [6,7,10,12]. As an example, a numerical grading scale (e.g., 1 to 5) can be used for assessment of group members [6,7,12], whereby "1" indicates "*didn't contribute*", "2" indicates "*willing but not successful*", "3" indicates "*average*", "4" indicates "*above average*", and "5" indicates "*outstanding*" [6,7,12]. Even though the use of numerals in peer assessment is popular, it suffers from problems associated with

* Corresponding author. Tel.: +60 146884403.
E-mail addresses: kcchai@live.com (K.C. Chai), kmtay@feng.unimas.my (K.M. Tay), chee.lim@deakin.edu.au (C.P. Lim).

psychological measurement in terms of the meaning pertaining to the numerals used (see [17] for a study on the theoretical relationship between measurement and marking). It would be more natural to define assessment grades using subjective and vague linguistic terms [17]. Furthermore, the conventional method aggregates individual scores to produce a total score. In some situations (as illustrated in Section 4.3), it is difficult to distinguish the ranking order of students using the same numeral score.

Fuzzy set theory has been used in education assessment [9,18–23]. It is useful to deal with linguistic grades such as "*didn't contribute*", "*willing but not successful*", and "*average*" in a grading system, which involve a substantial amount of fuzziness and vagueness [18]. It is worth mentioning that fuzzy set theory is an efficient and effective method to represent uncertainties [18,19]. Comparing with methods based on numerical grading scores [6,7,12], fuzzy set theory offers an alternative to linguistic evaluation in which "*fuzzy*" words, instead of numerals, are used during the assessment procedure [9,18–23]. Besides that, "*computing with words*", as coined by Zadeh, is also a methodology related to fuzzy set theory, whereby the objects of computation are words and propositions drawn from a natural language [24,25].

Motivated by the success of fuzzy set theory in education assessment [9,18–23], this paper aims to propose a fuzzy peer assessment methodology that evaluates each student's contribution in a group work. The proposed methodology is a synthesis of perceptual computing (Per-C) [25–29] and a fuzzy ranking algorithm that uses fuzzy preference relations [30]. The rationale for the proposed methodology hinges on a number of imperatives. Firstly, the available information is too imprecise to be justified with numerals, which is more suitable to be represented using words [25]. In this paper, Per-C is adopted owing to its effectiveness in handling inherent uncertainties in words [25]. Specifically, Per-C is able to handle subjectivity, vagueness, imprecision, and uncertainty while achieving tractability and robustness in modelling human decision-making behaviours [25–29]. Comparing with type-1 fuzzy models [9,18–23], Per-C adopts interval type-2 fuzzy sets (IT2FSs) in tackling a decision making problem [25–29]. IT2FS has more flexibility in preserving and processing uncertainties than type-1 fuzzy set [25]. Indeed, Per-C has been successfully implemented to undertake a number of fuzzy multiple criteria hierarchical decision making problems [25–29]. In [25–29], Per-C focuses on ranking the sequence of outcomes, mapping the outcomes into words and/or classifying the outcomes into different categories. Nevertheless, the use of Per-C in peer assessment is still new. In this paper, the relative importance of the outcomes, i.e., the contribution of each student with respect to those from other students, is examined in detail, which is yet to be investigated in the literature, e.g. [25–29]. In this aspect, our preliminary work [30], as discussed in Section 2.3, is further extended to serve this purpose. Then, the effectiveness and practicality of the proposed methodology are evaluated with a case study in an engineering course (i.e., Multiprocessors Architecture) at Universiti Malaysia Sarawak. The results from the conventional and proposed methodologies are analyzed and discussed. In essence, this paper contributes to a new fuzzy peer assessment methodology in which human linguistic words are adopted in the entire assessment process. Besides that, the proposed methodology provides an insight pertaining to each individual's contribution; therefore providing personalized assessment in a cooperative learning environment.

The rest of this paper is organized as follows. In Section 2, the background of fuzzy sets, perceptual computing, fuzzy ranking algorithms and peer assessment in problem-based learning is presented. In Section 3, a new technique for fuzzy peer assessment is explained in detail. In Section 4, a case study is conducted to demonstrate the usefulness of the proposed fuzzy peer assessment
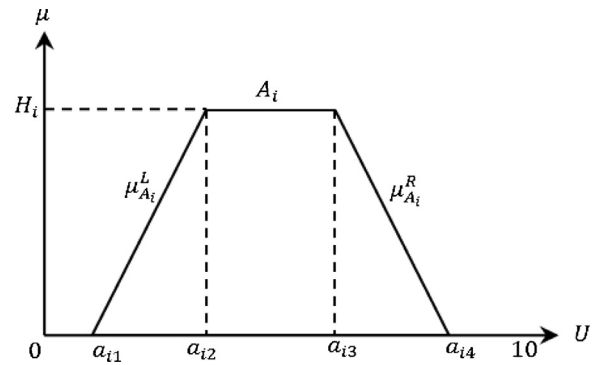

Fig. 1. The membership function of a trapezoidal fuzzy set.

methodology. Concluding remarks and suggestions for future research are presented in Section 5.

2. Preliminaries

A number of notations and definitions related to type-1 fuzzy sets (T1FSs) and interval type-2 fuzzy sets (IT2FSs) are presented in Section 2.1. A review on perceptual computing is presented in Section 2.2. Our preliminary work related to a fuzzy ranking algorithm is reviewed in Section 2.3. Finally, an overview on peer assessment in problem-based learning is presented in Section 2.4.

2.1. Definitions

Consider a set of trapezoidal T1FSs, i.e., A_i where $i = 1,2,\ldots,m$, in the universe of discourse, U. A trapezoidal T1FS A_i is parameterized as $A_i = (a_{i1}, a_{i2}, a_{i3}, a_{i4}, ; H_i)$, as illustrated in Fig. 1.

Definition 1. [30,31]: A fuzzy membership function, μA_i, of A_i, as shown in Fig. 1, is defined as follows:

$$\mu A_i(X) = \begin{cases} \mu_{A_i}^L(X), & a_{i1} \leq X \leq a_{i2}, \\ H_i, & a_{i2} \leq X \leq a_{i3}, \\ \mu_{A_i}^R(X), & a_{i3} \leq X \leq a_{i4}, \\ 0, & otherwise. \end{cases} \quad (1)$$

where $\mu_{A_i}^L$ is continuous and strictly increasing in interval $[a_{i1}, a_{i2}]$, as defined in Eq. (2), $\mu_{A_i}^R$ is continuous and strictly decreasing in interval $[a_{i3}, a_{i4}]$, as defined in Eq. (3), and $H_i \in [0,1]$. Besides that, $a_{i1}, a_{i2}, a_{i3}, a_{i4}$ are real values, i.e., $\forall a_{i1}, a_{i2}, a_{i3}, a_{i4} \in x$, such that $a_{i1} \leq a_{i2} \leq a_{i3} \leq a_{i4}$, and $\exists x \in U$.

$$\mu_{A_i}^L : [a_{i1}, a_{i2}] \rightarrow [(x - a_{i1})/(a_{i2} - a_{i1})]H_i \quad (2)$$

$$\mu_{A_i}^R : [a_{i3}, a_{i4}] \rightarrow [(a_{i4} - x)/(a_{i4} - a_{i3})]H_i. \quad (3)$$

Definition 2. [30,32,33]: An IT2FS $\tilde{A}_i$ is denoted as $\tilde{A}_i = (\bar{A}_i, \underline{A}_i)$, and $\tilde{A}_i$ is parameterized in Eq. (4). The upper and lower membership functions of $\tilde{A}_i$ (i.e., $\mu_{\bar{A}_i}$ and $\mu_{\underline{A}_i}$ respectively) are represented by type-1 membership functions.

$$\tilde{A}_i = (\bar{A}_i, \underline{A}_i) = ((\bar{a}_{i1}, \bar{a}_{i2}, \bar{a}_{i3}, \bar{a}_{i4}; \bar{H}_i), (\underline{a}_{i1}, \underline{a}_{i2}, \underline{a}_{i3}, \underline{a}_{i4}; \underline{H}_i)) \quad (4)$$

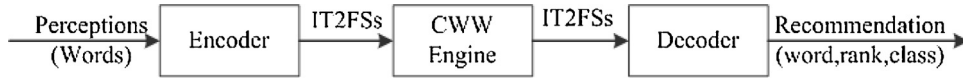**Definition 3.** [32]: The fuzzy addition operation between two IT2FSs is defined as follows:

Fig. 2. The structure of Per-C [25–29].

$\tilde{A}_1 \oplus \tilde{A}_2$

$= ((\bar{a}_{11}, \bar{a}_{12}, \bar{a}_{13}, \bar{a}_{14}; H_1), (\underline{a}_{11}, \underline{a}_{12}, \underline{a}_{13}, \underline{a}_{14}; \underline{H}_1)),$

$\qquad \oplus \left(\left(\bar{a}_{21}, \bar{a}_{22}, \bar{a}_{23}, \bar{a}_{24}; \bar{H}_2 \right), (\underline{a}_{21}, \underline{a}_{22}, \underline{a}_{23}, \underline{a}_{24}; \underline{H}_2) \right)$ (5)

$= \left(\left(\bar{a}_{11} + \bar{a}_{21}, \bar{a}_{12} + \bar{a}_{22}, \bar{a}_{13} + \bar{a}_{23}, \bar{a}_{14} + \bar{a}_{24}; \min\left(\bar{H}_1, \bar{H}_2 \right) \right), (\underline{a}_{11} + \underline{a}_{21}, \underline{a}_{12} +$

$\qquad \underline{a}_{22}, \underline{a}_{13} + \underline{a}_{23}, \underline{a}_{14} + \underline{a}_{24}; \min(\underline{H}_1, \underline{H}_2).$

Definition 4. [32]: The fuzzy multiplication operation between an IT2FS and crisp value k is defined as follows:

$k\tilde{A}_1$

$= ((k \times \bar{a}_{11}, k \times \bar{a}_{12}, k \times \bar{a}_{13}, k \times \bar{a}_{14}, k; \bar{H}_1),$ (6)

$\qquad (k \times \underline{a}_{11}, k \times \underline{a}_{12}, k \times \underline{a}_{13}, k \times \underline{a}_{14}, k; \underline{H}_1).$

2.2. Review on perceptual computing

The general structure of Per-C is depicted in Fig. 2. It consists of three components [25–29], i.e., an encoder, a computing-with-words (CWW) engine, and a decoder. Linguistic grades or words from humans are converted into IT2FSs through the encoder. The CWW engine aggregates the outputs from the encoder. The decoder maps the outputs of the CWW engine into a recommendation, which can be in the form of a word, rank, or class. However, the decoder ranks the outputs independently without considering their relative importance with respect to the recommendation. Nevertheless, it is imperative to rank each student's contribution and derive a set of performance indices that reflects the student's relative contribution in fuzzy peer assessment (i.e., peer assessment scores (*PA*)). While many fuzzy ranking algorithms are available in the literature, to the best of our knowledge, only a few solutions (e.g., [30] and [32]) are focused on the relative importance of two or more IT2FSs. The details are presented in Section 2.3.

2.3. Review on fuzzy ranking algorithms

As discussed earlier, our focus is on investigating the relative importance of two or more IT2FSs. As such, the methods in [30] and [32] are considered. Specifically, our preliminary work [30], which is an extension of that in [32], forms the foundation of the proposed method. In [30], we studied the rationality, i.e., the fulfilment of the six reasonable ordering properties as stated in [25,34], and presented a number of improvements as compared with those

Fig. 3. Example of two IT2FSs.

in [32]. Assume that a set of IT2FSs, $\tilde{A}_i$, where $i = 1, 2, \ldots, m$. To better clarify the explanation, a simulated example with two IT2FSs (i.e., $\tilde{A}_1 = ((5.03, 7.03, 8.25, 9.38; 1), (6.03, 7.75, 7.93, 8.82; 0.75))$ and $\tilde{A}_2 = ((4.03, 6.03, 7.07, 8.93; 1), (5.03, 6.57, 6.60, 8.00; 0.75))$, as illustrated in Fig. 3, is considered. The fuzzy ranking algorithm in [30] is summarized in five steps, as follows.

Step 1. Discretize the support of $\bar{A}_i$ and $\underline{A}_i$ of $\tilde{A}_i$ into N points, i.e., $x_{\bar{A}_i, k}$ and $x_{\underline{A}_i, k}$, $k = 1, 2, 3, \ldots, N$, and obtain $\mu_{\bar{A}_i}(x_{\bar{A}_i, k})$ and $\mu_{\underline{A}_i}(x_{\underline{A}_i, k})$, respectively.

As an example, the discretized points of $\bar{A}_i$ are expressed in a sequence of $x_{\bar{A}_i, 1}, x_{\bar{A}_i, 2}, \ldots, x_{\bar{A}_i, N}$, where $X\bar{A}_i, 1$ and $X\bar{A}_i, N$ are the left and right-end points of $\bar{A}_i$, respectively. The discretized points in the horizontal component of $\bar{A}_i$ and $\underline{A}_i$, i.e., $x_{\bar{A}_i, k}$ and $x_{\underline{A}_i, k}$ are computed using Eqs. (7) and (8), respectively. On the other hand, the discretized points in the vertical component of $\bar{A}_i$ and $\underline{A}_i$, i.e., $\mu_{\bar{A}_i}(x_{\bar{A}_i})$ and $\mu_{\underline{A}_i}(x_{\underline{A}_i})$ are computed using Eq. (1). The discretized points of $\bar{A}_i$ are expressed in Eqs. (9) and (11) as well as $\underline{A}_i$ are expressed in Eqs. (10) and (12), as follows.

$x_{\bar{A}_i, k} = x_{\bar{A}_i, 1} + (k - 1) \left(\frac{x_{\bar{A}_i, N} - x_{\bar{A}_i, 1}}{N - 1} \right).$ (7)

$x_{\underline{A}_i, k} = x_{\underline{A}_i, 1} + (k - 1) \left(\frac{x_{\underline{A}_i, N} - x_{\underline{A}_i, 1}}{N - 1} \right).$ (8)

$x_{\bar{A}_i} = \begin{bmatrix} x_{\bar{A}_1, 1} & \cdots & x_{\bar{A}_1, N} \\ \vdots & \ddots & \vdots \\ x_{\bar{A}_m, 1} & \cdots & x_{\bar{A}_m, N} \end{bmatrix}.$ (9)

$x_{\underline{A}_i} = \begin{bmatrix} x_{\underline{A}_1, 1} & \cdots & x_{\underline{A}_1, N} \\ \vdots & \ddots & \vdots \\ x_{\underline{A}_m, 1} & \cdots & x_{\underline{A}_m, N} \end{bmatrix}.$ (10)

$\mu_{\bar{A}_i}(x_{\bar{A}_i}) = \begin{bmatrix} \mu_{\bar{A}_1}(x_{\bar{A}_1, 1}) & \cdots & \mu_{\bar{A}_1}(x_{\bar{A}_1, N}) \\ \vdots & \ddots & \vdots \\ \mu_{\bar{A}_m}(x_{\bar{A}_m, 1}) & \cdots & \mu_{\bar{A}_m}(x_{\bar{A}_m, N}) \end{bmatrix}.$ (11)

$\mu_{\underline{A}_i}(x_{\underline{A}_i}) = \begin{bmatrix} \mu_{\underline{A}_1}(x_{\underline{A}_1, 1}) & \cdots & \mu_{\underline{A}_1}(x_{\underline{A}_1, N}) \\ \vdots & \ddots & \vdots \\ \mu_{\underline{A}_m}(x_{\underline{A}_m, 1}) & \cdots & \mu_{\underline{A}_m}(x_{\underline{A}_m, N}) \end{bmatrix}.$ (12)

With the example in Fig. 3, the left- and right- end points of $\bar{A}_1$ (i.e., 5.03 and 9.38) and $\bar{A}_2$ (i.e., 4.03 and 8.93) are discretized into $N = 1000$ using Eq. (7). The left- and right- end points of $\underline{A}_1$ and $\underline{A}_2$ are computed using Eq. (8). The discretized points for $\bar{A}_1$ and $\bar{A}_2$ as well as $\underline{A}_1$ and $\underline{A}_2$ are represented as matrices, i.e., Eqs. (9) and (11) as well as Eqs. (10) and (12), respectively.

Step 2. Compute $P(\bar{A}_i \geq \bar{A}_j)$ and $P(\underline{A}_i \geq \underline{A}_j)$ using Eqs. (13) and (14), respectively. Note that $P(\bar{A}_i \geq \bar{A}_j)$ and $P(\underline{A}_i \geq \underline{A}_j)$ are the ratios of the distance between two FSs of favourable outcomes to the total distance pertaining to two FSs of the entire possible outcomes.

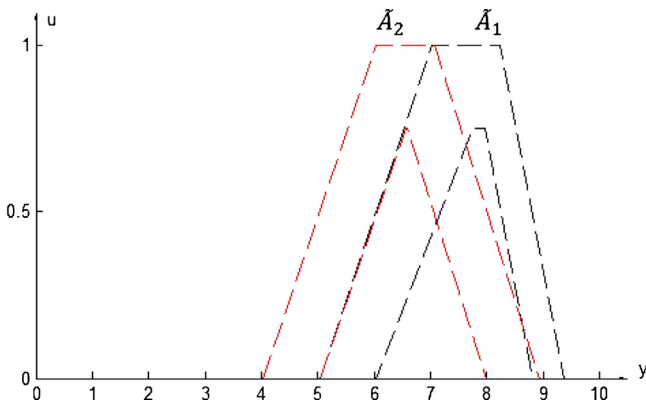$P(\bar{A}_i \geq \bar{A}_j) = \max \left(1 - \max \left(\frac{E_{ji}^{\bar{U}}}{E_{ji}^{\bar{L}}} \right), 0 \right).$ (13)

where,

$$\overline{E^U_{ji}} = \max_{k=1,2...,N}(x_{\bar{A}_j,k}) - \min_{k=1,2...,N}(x_{\bar{A}_i,k})$$

$$+ \frac{1}{N}\sum_{k=1}^{N}(\max(x_{\bar{A}_j,k} - x_{\bar{A}_i,k}, 0) + \max(\mu_{\bar{A}_j}(x_{\bar{A}_j,k}) - \mu_{\bar{A}_i}(x_{\bar{A}_i,k}), 0))$$

$$\overline{E^L_{ji}} = (x_{\bar{A}_i,N} - x_{\bar{A}_i,1}) + (x_{\bar{A}_j,N} - x_{\bar{A}_j,1})$$

$$+ \frac{1}{N}\sum_{k=1}^{N}\left(\left|x_{\bar{A}_j,k} - x_{\bar{A}_i,k}\right| + \left|\mu_{\bar{A}_j}(x_{\bar{A}_j,k}) - \mu_{\bar{A}_i}(x_{\bar{A}_i,k})\right|\right)$$

$i, j = 1, 2, 3, \ldots, m$

$$P(\underline{A}_i \geq \underline{A}_j) = \max\left(1 - \max\left(\frac{E^U_{ji}}{\underline{E^L_{ji}}}, 0\right), 0\right). \tag{14}$$

where,

$$\underline{E^U_{ji}} = \max_{k=1,2...,N}(x_{\underline{A}_j,k}) - \min_{k=1,2...,N}(x_{\underline{A}_i,k})$$

$$+ \frac{1}{N}\sum_{k=1}^{N}(\max(x_{\underline{A}_j,k} - x_{\underline{A}_i,k}, 0) + \max(\mu_{\underline{A}_j}(x_{\underline{A}_j,k}) - \mu_{\underline{A}_i}(x_{\underline{A}_i,k}), 0))$$

$$\underline{E^L_{ji}} = (x_{\underline{A}_i,N} - x_{\underline{A}_i,1}) + (x_{\underline{A}_j,N} - x_{\underline{A}_j,1})$$

$$+ \frac{1}{N}\sum_{k=1}^{N}(\left|x_{\underline{A}_j,k} - x_{\underline{A}_i,k}\right| + \left|\mu_{\underline{A}_j}(x_{\underline{A}_j,k}) - \mu_{\bar{A}_i}(x_{\bar{A}_i,k})\right|)$$
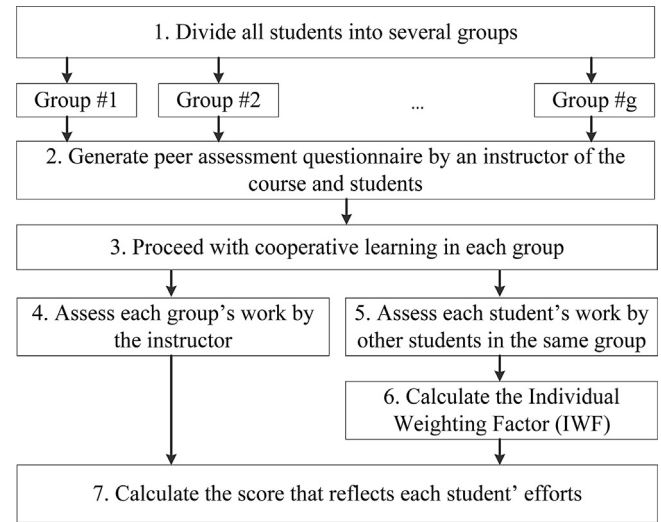
$i, j = 1, 2, 3, \ldots, m$

Step 3. Generate the fuzzy preference matrices as in Eqs. (15) and (16).

$$\bar{P} = \begin{bmatrix} p(\bar{A}_1 \geq \bar{A}_1) & \cdots & p(\bar{A}_1 \geq \bar{A}_j) & \cdots & p(\bar{A}_1 \geq \bar{A}_m) \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ p(\bar{A}_i \geq \bar{A}_1) & \cdots & p(\bar{A}_i \geq \bar{A}_j) & \cdots & p(\bar{A}_i \geq \bar{A}_m) \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ p(\bar{A}_m \geq \bar{A}_1) & \cdots & p(\bar{A}_m \geq \bar{A}_j) & \cdots & p(\bar{A}_m \geq \bar{A}_m) \end{bmatrix}. \tag{15}$$

$$\underline{P} = \begin{bmatrix} p(\underline{A}_1 \geq \underline{A}_1) & \cdots & p(\underline{A}_1 \geq \underline{A}_j) & \cdots & p(\underline{A}_1 \geq \underline{A}_m) \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ p(\underline{A}_i \geq \underline{A}_1) & \cdots & p(\underline{A}_i \geq \underline{A}_j) & \cdots & p(\underline{A}_i \geq \underline{A}_m) \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ p(\underline{A}_m \geq \underline{A}_1) & \cdots & p(\underline{A}_m \geq \underline{A}_j) & \cdots & p(\underline{A}_m \geq \underline{A}_m) \end{bmatrix}. \tag{16}$$

Based on the aforementioned example, the fuzzy preference matrices for $\bar{A}_1$ and $\bar{A}_2$ (i.e., $\bar{P}$) and for $\underline{A}_1$ and $\underline{A}_2$ (i.e., $\underline{P}$) are computed using Eq. (13) and Eq. (14) in Step 2, respectively. The simulation outcomes from Step 2 are compiled and written in Eqs. (15) and (16), as follows.

$$\bar{P} = \begin{bmatrix} 0.5000 & 0.6098 \\ 0.3902 & 0.5000 \end{bmatrix}, \underline{P} = \begin{bmatrix} 0.5000 & 0.7036 \\ 0.2964 & 0.5000 \end{bmatrix}$$

Fig. 4. The procedure to assess an individual student's contribution in a group project [9].

The flow chart:
1. Divide all students into several groups
Group #1   Group #2   ...   Group #g
2. Generate peer assessment questionnaire by an instructor of the course and students
3. Proceed with cooperative learning in each group
4. Assess each group's work by the instructor
5. Assess each student's work by other students in the same group
6. Calculate the Individual Weighting Factor (IWF)
7. Calculate the score that reflects each student' efforts

Step 4. Compute the ranking indices of $\bar{A}_i$ and $\underline{A}_i$ using Eqs. (17) and (18), respectively.

$$RI(\bar{A}_i) = \frac{1}{m(m-1)}\left(\sum_{j=1}^{m}p(\bar{A}_i \geq \bar{A}_j) + \frac{m}{2} - 1\right) \tag{17}$$

$$RI(\underline{A}_i) = \frac{1}{m(m-1)}\left(\sum_{j=1}^{m}p(\underline{A}_i \geq \underline{A}_j) + \frac{m}{2} - 1\right) \tag{18}$$

Again, based on the same example, using Eq. (17), the ranking indices for $\bar{A}_1$ and $\bar{A}_2$ are 0.5549 and 0.4451, respectively. Using Eq. (18), the ranking indices for $\underline{A}_1$ and $\underline{A}_2$ are 0.6018 and 0.3982, respectively.

Step 5. The ranking index of $\tilde{A}_i$ is computed using Eq. (19).

$$RI(\tilde{A}_i) = \frac{RI(\underline{A}_i) + RI(\bar{A}_i)}{2} \tag{19}$$

A higher ranking index of $\tilde{A}_i$ indicates a higher ranking of $\tilde{A}_i$. With the same example, using Eq. (19), the ranking indices of $\tilde{A}_1$ and $\tilde{A}_2$ are 0.5784 and 0.4216, respectively. Therefore, $\tilde{A}_1$ has a higher ranking than $\tilde{A}_2$.

2.4. Overview on peer assessment in problem-based learning

2.4.1. Background

A summative assessment procedure, as discussed in [9], is presented in Fig. 4. Firstly, students are divided into several groups, and a set of peer assessment questionnaires that has been agreed upon by the instructor and students is distributed. Then, each group proceeds with the process of cooperative learning. At the end of the cooperative learning process, the students are required to submit and present their findings for assessment. Their achievements are evaluated by the instructor as well as by their peers, as shown in the 5th stage in Fig. 4. The results from peer assessment (denoted as the *PA* scores) are aggregated and the Individual Weighting Factor (*IWF*) is derived, as in the 6th stage. *IWF* is a measure of each individual's contribution in his/her group. There are several methods to compute *IWF* [5,6]. In this paper, the method in [5] is adopted owing to a number of advantages, i.e., it discourages free-riders, encourages above-average contributions, and discourages individualistic behaviours in a team. Finally, the result that reflects each student's effort is computed.

Table 1
An example of peer assessment (i.e., group #1) using the conventional method [5,6].

Student Being Assessed		#1A			#1D		
Specific Criteria, i	Evaluators	#1B	#1C	#1D	#1A	#1B	#1C
	Weighing						
1. Participated in group meetings	(20%)	0.80	1.00	1.00	0.60	0.80	1.00
2. Communicated constructively to discussion	(20%)	1.00	1.00	1.00	1.00	0.80	1.00
3. Generally was cooperative in group activities	(15%)	0.60	0.60	0.60	0.60	0.75	0.60
4. Contributed to good problem-solving skills	(15%)	0.60	0.45	0.45	0.60	0.60	0.45
5. Contributed useful ideas	(10%)	0.50	0.30	0.30	0.50	0.40	0.30
6. Demonstrated good interest to task given	(10%)	0.40	0.40	0.50	0.40	0.50	0.50
7. Prepared drafts of report in good quality	(10%)	0.50	0.40	0.40	0.50	0.50	0.40
Aggregated Results from each evaluator		**4.40**	**4.15**	**4.25**	**4.20**	**4.35**	**4.25**

Table 2
An example of group assessment from the instructor.

Project Marking Form		
Group ID: #1		
General Aspects	Specific Criteria	Group Score
Project Report (60%)	1. Correct use of methods (25%)	20
	2. Design, simulate and analyze the system to meet the requirements of the project (25%)	23
	3. Delivery the outcomes of the project in written forms (25%)	19
	4. Structure and presentation of the project report (25%)	22
	Total mark (out of 60%):	**50.4**
Project Presentation (40%)	5. Delivery the outcomes of the project in oral forms (25%)	18
	6. Reliability of the proposed method (25%)	22
	7. Efficiency and flexibility of the proposed method (25%)	22
	8. Responses to questions and defend their projects with valid reasoning and arguments (25%)	20
	Total mark (out of 40%):	**32.8**
	Total group score (out of 100%)	**83.2**

In this paper, a cooperative learning activity pertaining to an engineering course (i.e., Multiprocessors Architecture) at Universiti Malaysia Sarawak, was studied. The goals of cooperative learning were to improve students' problem-solving skills in groups. Students were required to focus on a task related to computing problems, i.e., designing a multiprocessor-based system for data analysis.

A total of 28 students participated in the cooperative learning activity. In accordance with Fig. 4, they were divided into 7 groups by the course instructor. Each group consists of four members with different proficiency levels and background. The students were required to complete their project within a given deadline. Then, they were required to submit their project report and present the software that they developed. Peer assessment was conducted using an evaluation form provided by the instructor. Each student was assessed based on seven criteria, as shown in Table 1. Each criterion was tagged with a weight predefined by the instructor, i.e., W_1, W_2, W_3, W_4, W_5, W_6, and W_7, corresponding to 20%, 20%, 15%, 15%, 10%, 10%, and 10%, respectively. An example of the students' peer assessment results is shown in Table 1.

Each group (i.e., #g, where g = 1, 2, 3, ..., 7) consisted of four students (i.e., $S_{g,s} \in [S_{g,A}, S_{g,B}, S_{g,C}, S_{g,D}]$). $S_{g,R}$ and $S_{g,r}$ represented the student being assessed and the (three) evaluators, where $S_{g,R}$, $S_{g,r} \in S_{g,s}$, respectively. Self-evaluation was not practiced, i.e., $S_{g,R} \notin S_{g,r}$. $W_{g,r}$ indicated the evaluator's weight (i.e., a measure of the evaluator's expertise level). In this study, each evaluator was given an equal weight. As an example of the peer assessment structure of group #1 is depicted in Fig. 5.

Besides that, each group was assessed by the instructor based on two general criteria i.e., project report and project presentation, as shown in Table 2. These two general criteria were further divided into a number of specific criteria. An example of the group assessment structure (i.e., for group #1) is shown in Table 2, whereby group #1 scored 83.2% out of 100%.

2.4.2. The conventional peer assessment methodology

The conventional peer assessment methodology in [5,6] is reviewed. The methodologies, as presented in [5,6], can be represented by Eqs. (20) and (21). Firstly, a peer assessment score (i.e., *PA*), which is the ratio of the actual sum of scores to the highest possible score, is calculated. Then, the *IWF* that reflects each student's contribution is computed using Eq. (21) [5]. Finally, each individual's mark is computed using the parabolic equation of Eq. (22), where α is the scale factor that has an impact on the spread of the individual marks. In this paper, α = 1.2 is adopted (as indicated in [5]).

$$PA = \frac{Actual\ sum\ scored}{Highest\ possible\ score} \tag{20}$$

$$IWF = \frac{PA}{Average\ PA\ score} \tag{21}$$

$$IM = GM \times \begin{cases} IWF, & IWF \leq 1 \\ IWF - \dfrac{(IWF-1)^2}{2\beta}, & 1 < IWF < 1+\beta \\ 1 + 0.5\beta, & IMF \geq 1+\beta \end{cases} \tag{22}$$

where $\beta = \alpha(1 - 0.01\,GM)$, *GM* denotes the group mark from the instructor

Conventionally, the five-point Likert scale is commonly used, whereby "1", "2", "3", "4", and "5" indicate "*poor*", "*below average*", "*average*", "*above average*", and "*excellent*", respectively. Consider the example shown in Table 1. Students #1A and #1D are evaluated. Student #1A is rated 4, 5, and 5 by Evaluators #1B, #1C, and #1D, respectively, for the first specific criterion i.e., "participated in group meetings". Since the criterion is weighted at 20%, the ratings (i.e., 4, 5, and 5) are normalized to 0.80, 1.00, and 1.00, respectively, as presented in Table 1. By applying Eqs. (20) and (21), *PA* and IWF are computed. The results are tabulated in Table 3.
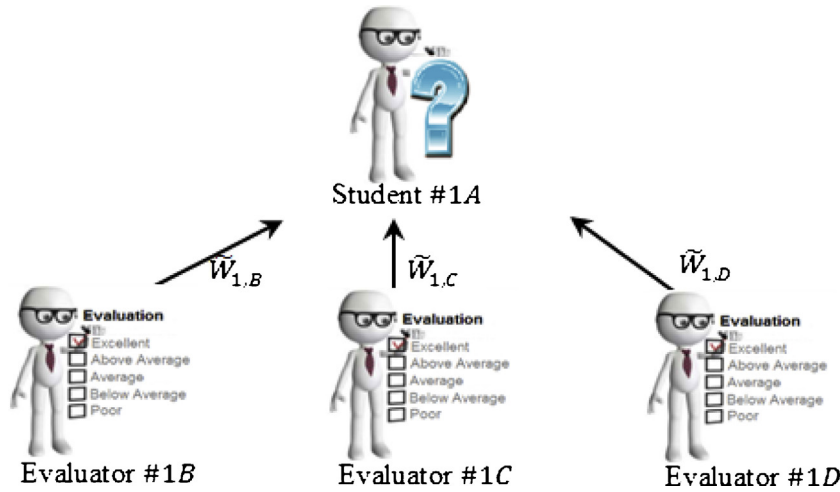
Fig. 5. An example of peer assessments in a group # 1 for assessing student #1A.

Table 3
The outcomes of Group #1 using the conventional methodology.

Students	#1A	#1D
Peer assessment score, PA	0.8533	0.8533
Individual Weighting Factor, IWF	1.0503	1.503

3. A new methodology for fuzzy peer assessment

In this section, the proposed new methodology for fuzzy peer assessment is described. Fig. 6 shows the overall fuzzy peer assessment framework. In this paper, the focus is on enhancing the 5th and 6th stages of the conventional methodology (as illustrated in Fig. 4) by using the Per-C paradigm [25–29]. The details of the proposed methodology are summarized in stages 5.1–5.4 and stage 6.0, as follows.

3.1. Encoder

The encoder converts human linguistic words into IT2FSs. This operation is performed based on the Interval Approach (IA),

whereby the IA codebook can be found in [25–27]. In this paper, five words (i.e., *Poor*, *Below Average*, *Average*, *Above Average*, and *Excellent*, abbreviated as P, BA, A, AA, and EX) are retrieved. The details of these five words (namely linguistic grades) are summarized in Table 4 and illustrated in Fig. 7. Each linguistic grade pertaining to the specific criterion is further described by the instructor. The interpretation of these linguistic grades is explained in Table A1 (Appendix A), which provides the details to each evaluator to aid his/her assessments.

3.2. Computing-with-words

Table 5 depicts the assessment of students #1A and #1D with linguistic grades. Using the CWW engine, the assessments from all evaluators are firstly aggregated with Eq. (23).

$$\tilde{y}_{g,r} = \sum_{i=1}^{7} \tilde{X}_{g,r,i} w_i. \tag{23}$$

where $\tilde{X}_{g,r,i}$ denotes the linguistic grade given by an evaluator (i.e., r), g indicates the group that r belongs to (i.e., g = 1, 2, 3, . . ., 7), and
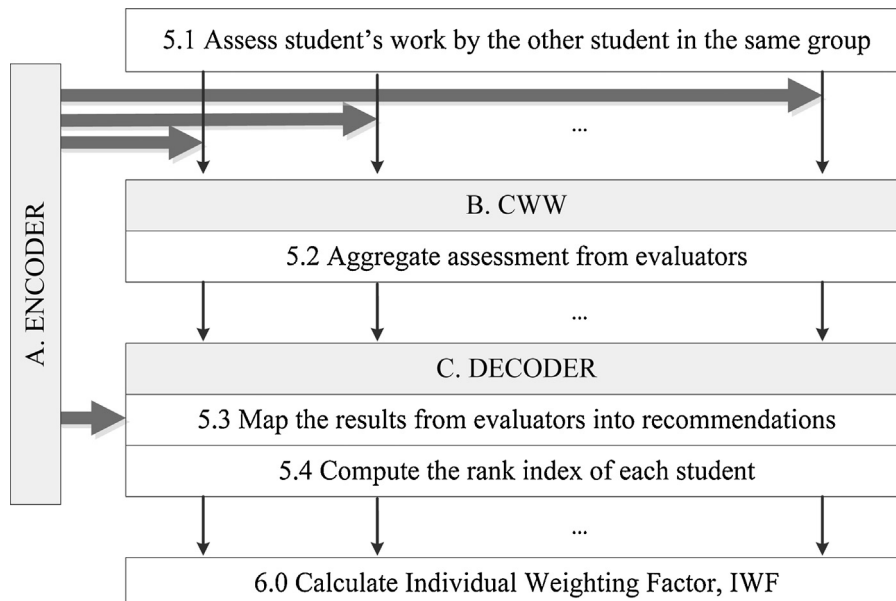

Fig. 6. The proposed fuzzy peer assessment framework.

Table 4
IT2FS for linguistic terms [25,26].

Linguistic Grades	Abbreviations	UMF $(\bar{a}_1, \bar{a}_2, \bar{a}_3, \bar{a}_4; \bar{H})$	LMF $(\underline{a}_1, \underline{a}_2, \underline{a}_3, \underline{a}_4; \underline{H})$
'Poor'	P	[0.00, 0.00, 1.50, 3.50; 1]	[0.00, 0.00, 0.50, 2.50; 1.00]
'Below Average'	BA	[0.50, 2.50, 3.50, 5.50; 1]	[1.50, 3.00, 3.00, 4.50; 0.75]
'Average'	A	[2.50, 4.50, 5.50, 7.50; 1]	[3.50, 5.00, 5.00, 6.50; 0.75]
'Above Average'	AA	[4.50, 6.50, 7.50, 9.50; 1]	[5.50, 7.00, 7.00, 8.50; 0.75]
'Excellent'	E	[6.50, 8.50, 10.0, 10.0; 1]	[7.50, 9.50, 10.0, 10.0; 1.00]
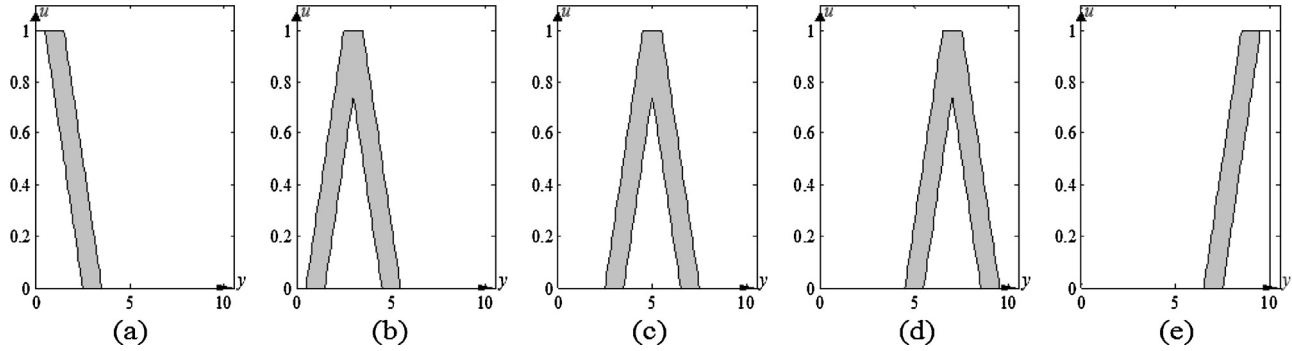
Fig. 7. Linguistic grades (a) *Poor*; (b) *Below Average*; (c) *Average*; (d) *Above Average* and (e) *Excellent*.

Table 5
An example of peer assessment using the proposed methodology.

Student Being Assessed		#1A			#1D		
Specific Criteria, i	Evaluators	#1B	#1C	#1D	#1A	#1B	#1C
	Weighing						
1. Participated in group meetings	(20%)	AA	EX	EX	A	AA	EX
2. Communicated constructively to discussion	(20%)	EX	EX	EX	EX	AA	EX
3. Generally was cooperative in group activities	(15%)	AA	AA	AA	AA	EX	AA
4. Contributed to good problem-solving skills	(15%)	AA	AA	A	AA	AA	A
5. Contributed useful ideas	(10%)	EX	EX	A	EX	AA	A
6. Demonstrated good interest to task given	(10%)	AA	AA	EX	AA	EX	EX
7. Prepared drafts of report in good quality	(10%)	EX	EX	AA	EX	EX	AA

i denotes the specific criterion (i.e., $i = 1, 2, 3, \ldots, 7$). As an example, $\tilde{X}_{1,B,1}$ is the linguistic grade given to the first specific criterion (i.e., "participated in group meetings") by Evaluator B in group #1 (i.e., #1B) and $\tilde{X}_{1,B,1}$ is AA in Table 5. Note that w_i denotes the normalized weight of W_i (i.e., $w_i \in [0, 1]$). As an example, $W_1 = 20\%$ and $w_1 = 0.20$. On the other hand, $\tilde{y}_{g,r}$ denotes the aggregated outcomes of Evaluator r in group #g.

The results from all evaluators are aggregated using Eq. (24).

$$\tilde{Y}_{g,R} = \frac{\sum_{r\in s, R \not\equiv r} \tilde{y}_{g,r} \tilde{W}_{g,r}}{\sum_{r\in s, R \not\equiv r} \tilde{W}_{g,r}}. \tag{24}$$

where $\tilde{W}_{g,r}$ indicates the expertise level (or weight) of Evaluator r in group #g. In this paper, $\tilde{W}_{g,A} = \tilde{W}_{g,B} = \tilde{W}_{g,C} = \tilde{W}_{g,D}$ i.e., all students are equally weighted. Therefore, Eq. (24) can be simplified to

$$\tilde{Y}_{g,R} = \frac{1}{|r|} \sum_{r\in s, R \not\equiv r} \tilde{y}_{g,r}. \tag{25}$$

where $|r|$ refers to the number of evaluators involved. Notice that fuzzy addition and multiplication operations (as in Definitions 3 and 4) are used in Eqs. (23), (24) and (25).

3.3. Decoder

The roles of the decoder are two-fold, i.e., to map the aggregated outcomes (represented in IT2FS, i.e., $\tilde{Y}_{g,R}$) from CWW to recommendation in term of words (Stage 5.3) and to rank the aggregated outcomes of students in the same group (Stage 5.4). The ranking indices are used to derive IWF in the latter stage. The former attempts to map the aggregated outcomes, i.e., $\tilde{Y}_{g,R}$, to the words listed in Table 4. The associated IT2FSs of these linguistic grades, i.e., *poor, below average, average, above average* and *excellent*, are denoted as $\tilde{Y}_P$, $\tilde{Y}_{BA}$, $\tilde{Y}_A$, $\tilde{Y}_{AA}$, and $\tilde{Y}_{EX}$, respectively, and their corresponding IT2FSs are shown in Fig. 7. Using the Jaccard similarity indicator [25], the similarity indicators between $\tilde{Y}_{g,R}$ and $\tilde{Y}_{LG}$ are denoted as $SJ(\tilde{Y}_{g,R}, \tilde{Y}_{LG})$, where $\tilde{Y}_{LG} \in [\tilde{Y}_P, \tilde{Y}_{BA}, \tilde{Y}_A, \tilde{Y}_{AA}, \tilde{Y}_{EX}]$. Eq. (26) shows the similarity indicator. $SJ(\tilde{Y}_{g,R}, \tilde{Y}_{LG}) \cong 1$ indicates $\tilde{Y}_{g,R}$ is similar to $\tilde{Y}_{LG}$.

$$SJ(\tilde{Y}_{g,R}, \tilde{Y}_{LG}) \cong \frac{\sum_{t=1}^{T} \min(\bar{\mu}_{\tilde{Y}_{g,R}}(y_t), \bar{\mu}_{\tilde{Y}_{LG}}(y_t)) + \sum_{t=1}^{T} \min(\underline{\mu}_{\tilde{Y}_{g,R}}(y_t), \underline{\mu}_{\tilde{Y}_{LG}}(y_t))}{\sum_{t=1}^{T} \max(\bar{\mu}_{\tilde{Y}_{g,R}}(y_t), \bar{\mu}_{\tilde{Y}_{LG}}(y_t)) + \sum_{t=1}^{T} \max(\underline{\mu}_{\tilde{Y}_{g,R}}(y_t), \underline{\mu}_{\tilde{Y}_{LG}}(y_t))}. \tag{26}$$

where $\bar{\mu}_{\tilde{Y}_{g,R}}(y_t)$ and $\underline{\mu}_{\tilde{Y}_{g,R}}(y_t)$ as well as $\bar{\mu}_{\tilde{Y}_{LG}}(y_t)$ and $\underline{\mu}_{\tilde{Y}_{LG}}(y_t)$ are the upper and lower membership functions of $\tilde{Y}_{g,R}$ and $\tilde{Y}_{LG}$, respectively. $y_t(t = 1, 2, \ldots, T)$ are equally spaced in the support region of $\tilde{Y}_{g,R} \cup \tilde{Y}_{LG}$. Consider an output space Y. $\tilde{Y}_{g,R}$, $\tilde{Y}_{LG}$, and y_t are the elements in Y (i.e., $\tilde{Y}_{g,R}, \tilde{Y}_{LG}, y_t \in Y$). In this paper, each aggregated outcome, i.e., $\tilde{Y}_{g,R}$, is mapped to two linguistic grades with the two largest similarity measures. $SJ(\tilde{Y}_{g,R}, \tilde{Y}_{LG})$ is represented

Table 6
Interpretation of each student's contribution.

Linguistic Grades	Poor, ($\tilde{Y}_P$)	Below Average, ($\tilde{Y}_{BA}$)	Average, ($\tilde{Y}_A$)	Above Average, ($\tilde{Y}_{AA}$)	Excellent, ($\tilde{Y}_{EX}$)
Interpretation	A 'sleeping' team member who relies on others to complete the overall project.	A passive team member who does not care about the overall project.	A good team member but needs to try harder to complete the overall project.	A strong group member who tries hard to complete the project.	A group leader who works hard to complete the overall project.

in percentage (%) by multiplying $SJ(\tilde{Y}_{g,R}, \tilde{Y}_{LG})$ with 100%. These similarity indicators map the outcomes, $\tilde{Y}_{g,R}$, into recommendations as described in Table 6.

To rank the aggregated outcomes of students in the same group, consider group #g consists of four students and the assessment outcomes are denoted as $\tilde{Y}_{g,R} \in [\tilde{Y}_{g,A}, \tilde{Y}_{g,B}, \tilde{Y}_{g,C}, \tilde{Y}_{g,D}]$. $\tilde{Y}_{g,R}$ are the fuzzy outcomes represented by IT2FSs. These fuzzy outcomes are ranked based on the algorithm proposed in [30] and reviewed in Section 2.3. The results are denoted as $RI(\tilde{Y}_{g,R}) \in [RI(\tilde{Y}_{g,A}), RI(\tilde{Y}_{g,B}), RI(\tilde{Y}_{g,C}), RI(\tilde{Y}_{g,D})]$ in which $RI(\tilde{Y}_{g,A}) + RI(\tilde{Y}_{g,B}) + RI(\tilde{Y}_{g,C}) + RI(\tilde{Y}_{g,D}) = 1$. Based on the ranking index, $RI(\tilde{Y}_{g,R})$, the IWF of each student (i.e., $IWF(\tilde{Y}_{g,R})$) is obtained using Eq. (27).

$$IWF(\tilde{Y}_{g,R}) = |S| \times RI(\tilde{Y}_{g,R}). \quad (27)$$

where $|S|$ denotes the number of students in the group.

Subsequently, each student's mark, IM, can be derived by substituting the results from Eq. (27) into Eq. (22).

4. A case study

In this section, a real case study is conducted. A variety of aspects from the initial survey to the end results are discussed and analyzed.

4.1. Survey

A survey on the selection of the type of assessment grades (i.e., linguistic grades or numerals) for peer assessment was firstly conducted. The responses indicated that the use of both linguistic grades and numeral was acceptable. While linguistic grades were perceived as more natural than numerals for peer assessment, a clear description for each linguistic grade should be provided. Therefore, a guideline, as presented in Table A1 (Appendix A), was prepared.

4.2. Results

The simulation results from the proposed fuzzy peer assessment methodology are represented as IT2FSs, as shown in Fig. 8. As an example, Fig. 8(a) presents the result of students in group #1. The legends (i.e., #1A, #1B, #1C and #1D) indicate the four students in group #1. The mark of student #1A is represented as an IT2FS. The detailed results from both conventional and proposed methodologies are presented in Table 7. Column "Students" indicates the student identification, columns "Conventional IWF" and "Proposed IWF" show the results from Eqs. (21) and (27), respectively. Column "GM" indicates the group marks from the instructor. Columns "IM with Conventional IWF (%)" and "IM with Proposed IWF (%)" indicate each student's marks, as the outcome of Eq. (22). Column "Peer Rating" indicates the results from Eq. (26). As an example, row "#1A" represents student A in group #1. He/she attains IMF = 1.0503 and IMF = 1.0869 using the conventional and proposed methodologies. His/her group is awarded an overall assessment of 83.20%. The conventional methodology indicates that student #1A's individual mark is 86.90%. Meanwhile, the proposed methodology awards student #1A 88.87%, and he/she is categorized as an "above average" (65.1%) and "excellent" (25.5%) student.

4.3. Analysis

The results in Table 7 are further illustrated in Fig. 9 for the purpose of visualization. Fig. 9 shows the group marks as well as the scores of each student based on the conventional and proposed methodologies. The green, blue, and red rectangular bars, indicate the group marks, IM (individual mark) using the conventional methodology, and IM using the proposed methodologies, respectively. The group marks do not reflect individual contributions as all students in the same group receive the same score. As an example, students #1A, #1B, #1C, and #1D attain the same score of 83.20%. Such assessment approach can be supplemented with the conventional and proposed methodologies. As can be seen in Fig. 9, the outcomes from the conventional and proposed methodologies are well-correlated; but the proposed methodology is able to distinguish students with similar scores. As an example, students #1A, #1B, #1C, and #1D are given 86.90%, 77.14%, 80.90%, and 86.90% respectively, using the conventional methodology; therefore they are ranked 1, 4, 3, and 1, respectively. Meanwhile, these students are given 88.87%, 72.79%, 79.01%, and 89.95%, respectively, using the proposed methodology; therefore they are ranked 2, 4, 3, and 1, respectively. The conventional methodology provides the same IM for students #1A and #1D (i.e., 86.90%). However, their performance can be distinguished with the proposed methodology.

It is also worth mentioning that the proposed methodology provides both crisp scores and recommendations. As an example, IT2FS for student #1A, as in Fig. 8(a), can be defuzzified to provide two meaningful interpretations i.e., a crisp score of 88.87% and a recommendation of the achievement being "above average" (65.1%) and "excellent" (25.5%). From Table 6, student #1A is most likely a strong group member who tries hard to complete the project ("above average"). Student #1A is also potentially a group leader who works hard to complete the project ("excellent"). Compared with the conventional methodology which only provides an individual mark, the proposed methodology provides more meaningful evaluation outcomes.

4.4. Similarity measure

The results from the fuzzy ranking algorithm [30] and Jaccard similarity measure algorithm [25] are in good agreement. As an example, $IWF(\tilde{Y}_{1,A})$ and $IWF(\tilde{Y}_{1,D})$ are 1.0869 and 1.0886, respectively. This shows that $IWF(\tilde{Y}_{1,D})$ is slightly larger than $IWF(\tilde{Y}_{1,A})$, which is in agreement with both students' scores, i.e., student #1A with AA (65.1%), EX (25.5%) and student #1D with AA (65.1%), EX (25.6%). The results of students #1A and #1D are indeed very close, but their contributions are slightly different.

4.5. Detection of possible free riders

Another key benefit of the proposed methodology is its ability to detect free rider(s) in a group. An illustrative example is as follows. Assume that student #1A is a free rider, and his/her grade is "poor", as shown in Table 8.

The aggregated results using the proposed methodology are illustrated in Fig. 10 and Table 9. The group marks do not reflect individual contributions as students from the same group are given
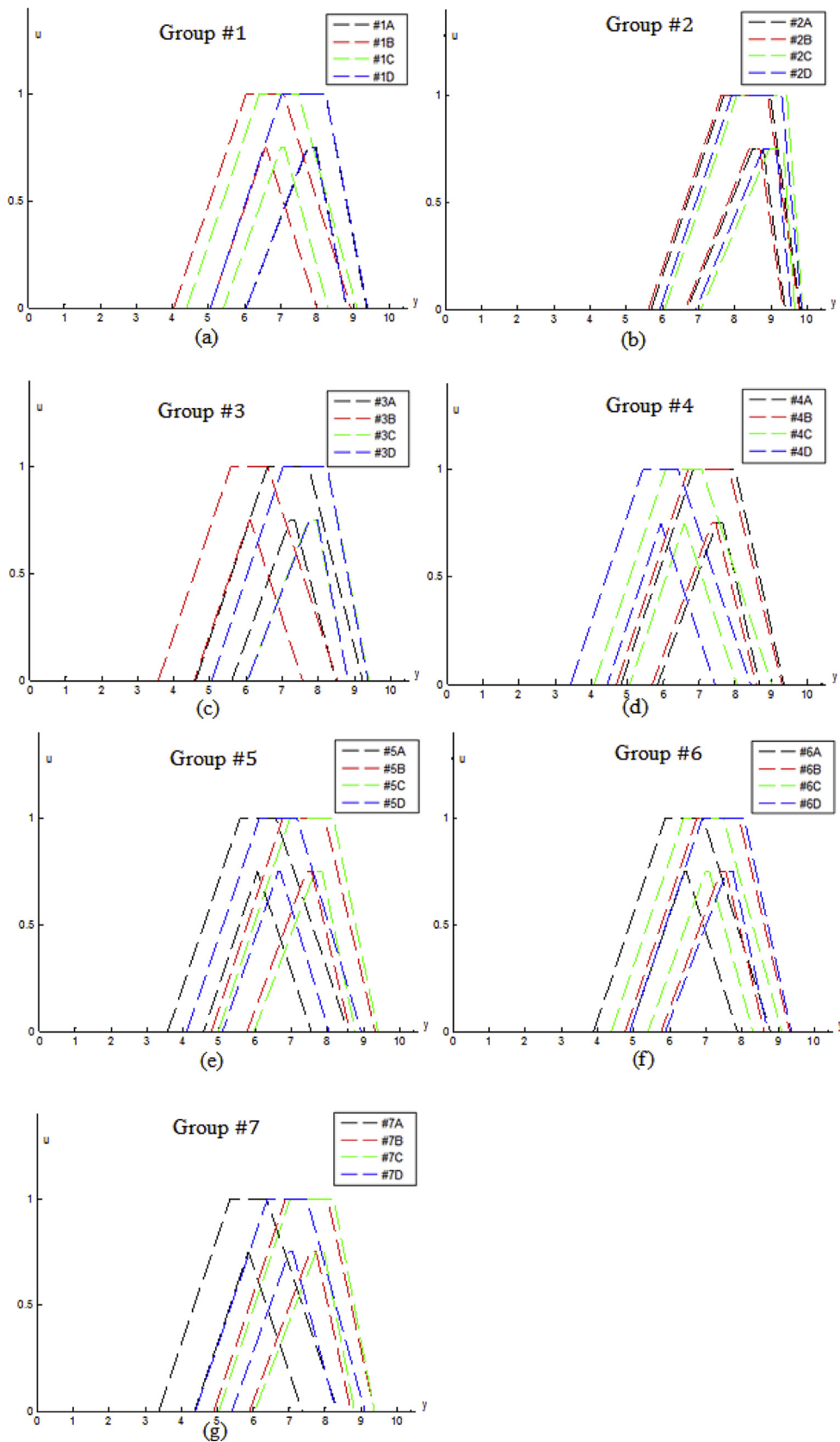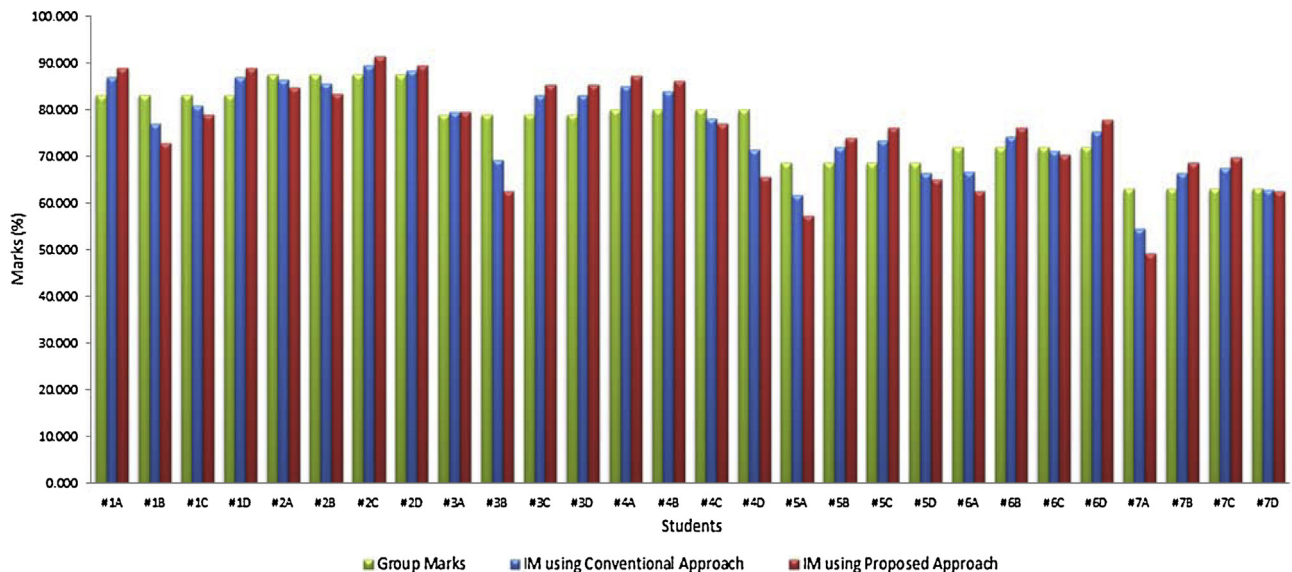
Fig. 8. Aggregated results of the students in each group: (a) Group #1; (b) Group #2; (c) Group #3; (d) Group #4; (e) Group #5; (f) Group #6; (g) Group #7.

Table 7
The outcomes using the conventional and proposed methodologies.

Students	Conventional *IWF*	Proposed *IWF*	GM (%)	*IM* with Conventional *IWF* (%)	*IM* with proposed *IWF* (%)	Peer Rating (Recommendations)
#1A	1.0503	1.0869	83.200	86.900	88.870	AA(65.1%), EX(25.5%)
#1B	0.9272	0.8749	83.200	77.141	72.790	A(27.4%), AA(67.4%)
#1C	0.9723	0.9497	83.200	80.896	79.013	A(19.8%), AA(90.7%)
#1D	1.0503	1.0886	83.200	86.900	88.951	AA(65.1%), EX(25.6%)
#2A	0.9857	0.9686	87.600	86.349	84.849	AA(37.6%), EX(48.3%)
#2B	0.9785	0.9533	87.600	85.723	83.513	AA(39.8%), EX(45.5%)
#2C	1.0250	1.0549	87.600	89.620	91.520	AA(27.6%), EX(68.0%)
#2D	1.0107	1.0232	87.600	88.507	89.474	AA(30.9%), EX(60.1%)
#3A	1.0052	1.0078	79.000	79.405	79.609	AA(16.5%), EX(89.3%)
#3B	0.8769	0.7922	79.000	69.278	62.587	A(41.1%), AA(45.7%)
#3C	1.0590	1.1008	79.000	83.154	85.372	AA(65.1%), EX(25.7%)
#3D	1.0590	1.0991	79.000	83.154	85.290	AA(65.1%), EX(25.5%)
#4A	1.0741	1.1215	80.000	85.084	87.259	AA(76.0%), EX(21.8%)
#4B	1.0569	1.0934	80.000	84.056	86.017	AA(83.8%), EX(19.4%)
#4C	0.9753	0.9636	80.000	78.024	77.086	A(26.6%), AA(69.5%)
#4D	0.8937	0.8216	80.000	71.493	65.725	A(46.0%), AA(41.1%)
#5A	0.9002	0.8338	68.600	61.755	57.200	A(41.1%), AA(46.0%)
#5B	1.0531	1.0889	68.600	72.004	73.980	AA(79.8%), EX(20.5%)
#5C	1.0786	1.1304	68.600	73.470	75.997	A(68.6%), EX(24.3%)
#5D	0.9682	0.9469	68.600	66.415	64.956	A(25.9%), AA(71.0%)
#6A	0.9260	0.8701	72.000	66.669	62.644	A(30.8%), AA(60.3%)
#6B	1.0344	1.0614	72.000	74.360	76.017	AA(79.8%), EX(20.5%)
#6C	0.9885	0.9778	72.000	71.174	70.404	A(19.8%), AA(90.7%)
#6D	1.0511	1.0907	72.000	75.421	77.650	AA(72.3%), EX(23.0%)
#7A	0.8665	0.7805	63.000	54.587	49.170	A(48.7%), AA(38.8%)
#7B	1.0599	1.1009	63.000	66.541	68.632	AA(72.3%), EX(23.0%)
#7C	1.0768	1.1262	63.000	67.450	69.820	AA(65.1%), EX(25.5%)
#7D	0.9969	0.9925	63.000	62.801	62.526	A(19.8%), AA(90.7%)

Table 8
An illustrative example i.e., student #1A is assessed as "poor" for the entire assessment.

Student Being Assessed			#1A		
Specific Criteria, *i*	Evaluators		#1B	#1C	#1D
	Weighing				
1. Participated in group meetings	(20%)		P	P	P
2. Communicated constructively to discussion	(20%)		P	P	P
3. Generally was cooperative in group activities	(15%)		P	P	P
4. Contributed to good problem-solving skills	(15%)		P	P	P
5. Contributed useful ideas	(10%)		P	P	P
6. Demonstrated good interest to task given	(10%)		P	P	P
7. Prepared drafts of report in good quality	(10%)		P	P	P

Fig. 9. Peer assessment results for 28 students using three different methods.
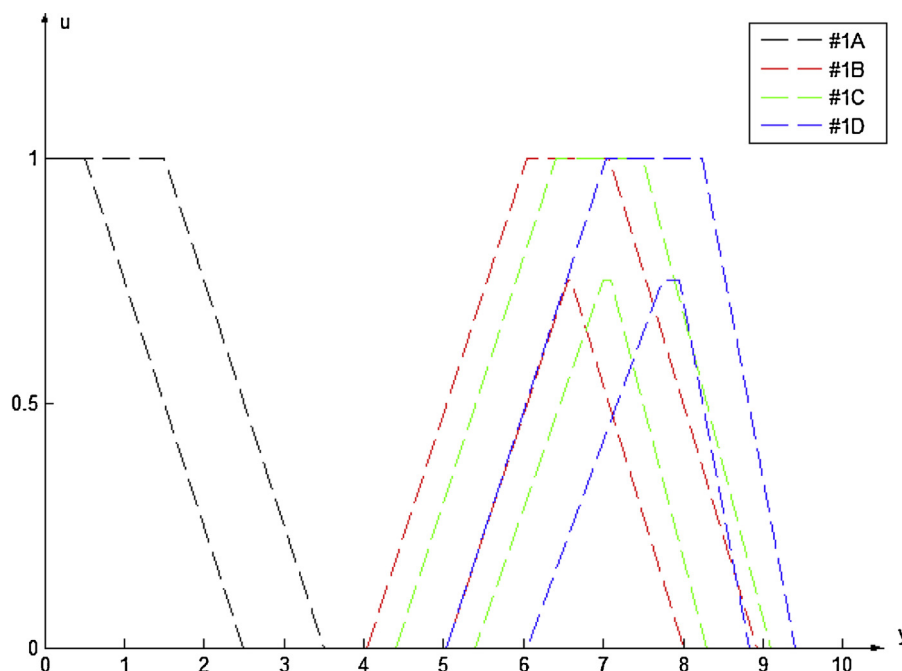
Fig. 10. Aggregated results of the students in Group #1 with possible free rider.

Table 9
The outcomes for the illustrative example.

Students	Conventional *IWF*	Proposed *IWF*	*GM* (%)	*IM* with Conventional *IWF* (%)	*IM* with proposed *IWF* (%)	Peer Rating (Recommendations)
#1*A*	0.3081	0.5000	83.200	25.633	41.600	P(100.0%)
#1*B*	1.1605	1.0938	83.200	91.646	89.186	A(27.4%), AA(67.4%)
#1*C*	1.2169	1.1515	83.200	92.285	91.069	A(19.8%), AA(90.7%)
#1*D*	1.3145	1.2548	83.200	92.285	91.587	AA(65.1%), Ex(25.6%)

the same score, i.e., 83.20%. The *IM* scores using the conventional methodology are 25.63%, 91.65%, 92.29% and 92.29% for students #1*A*, #1*B*, #1*C* and #1*D*, respectively. Even though student #1*A* obtains a low score, it is hard to ascertain the degree of the student being a free-rider. With the proposed methodology, the *IM* scores for students #1*A*, #1*B*, #1*C* and #1*D* are 41.60%, 90.59%, 91.59% and 91.59%, respectively. Student #1*A* is also assessed as "*P*(100%)". According to Table 6, student #1*A* is identified as a 'sleeping' team member who relies on others to complete the overall project, and the assessment outcome of "P(100%)" clearly indicates the degree of the student being a free rider in a group activity.

5. Conclusions

In this paper, a new fuzzy peer assessment methodology is proposed. The proposed methodology is the synthesis of Per-C and a fuzzy ranking algorithm that uses fuzzy preference relations. Per-C is able to provide assessment pertaining to students' contributions in term of recommendations in words. Meanwhile, the fuzzy ranking algorithm is able to provide indices that reflect students' contributions in a group, and subsequently rank them accordingly. These indices are employed to derive students' *IWF* scores. The group marks given by the instructor are then assigned to the students in the respective group based on their associated *IWF* scores. A real case study on an engineering course (i.e., Multiprocessors Architecture) has been conducted. The results from the conventional and proposed methodologies are compared, analyzed, and discussed in detail. The outcomes positively demonstrate that the proposed methodology is a potential solution to

undertake fuzzy peer assessment tasks, which involve vagueness and uncertainty, in cooperative learning environments. In short, the proposed approach has shown to be a useful potential solution to several shortcomings related to the use of numerals in peer assessment, e.g., the meaning of numerals, the difficulty to distinguish the ranking order of students, and the difficulty in interpreting peer assessment outcomes.

In this paper, a small population of students is used as a case study to demonstrate the viability of the proposed approach in peer assessment. As future works, the reliability of the proposed approach will be evaluated via a comprehensive study with larger size of population. Besides, consensus in group decision making [35,36] will be incorporated into the proposed methodology. The benefits include extending the proposed methodology from summative assessment to a hybrid framework of summative and formative assessments. Besides that, the proposed methodology will be deployed in other application domains such as risk assessment [37] and energy research [38].

Appendix A.

See Table A1.

Table A1
Peer Assessment form (Part 1: Details of Linguistic grades to each specific criterion).

Peer Assessment Form
Group ID: _____
Noted: Assessment of others is an important skill. You should take time to answer these questions, forcing yourself to be objective and unbiased. The individual response
will be kept confidential.
Please use the following linguistic grades to assess your friends. The interpretations of each linguistic grade are summarized as follows.
Linguistic grades: **Poor**, **Below Average**, **Average**, **Above Average** and **Excellent**
Scoring Rubric: Group Process

Specific Criteria	Linguistic grades and their respective interpretations				
	Poor	Below Average	Average	Above Average	Excellent
1. Participated in group meetings	Does not provide any ideas when participating in the group and in classroom discussion. Refuses to participate.	Rarely provides useful ideas when participating in the group and in classroom discussion. May refuse to participate.	Sometimes provides useful ideas when participating in the group and in classroom discussion. A satisfactory group member who does what is required.	Usually provides useful ideas when participating in the group and in classroom discussion.	Routinely provides useful ideas when participating in the group and in classroom discussion.
2. Communicated constructively to discussion	Does not listen to other team members.	Rarely listens to, shares with, and supports the efforts of others.	Often listens to, shares with, and supports the efforts of others	Usually listens to, shares, with, and supports the efforts of others.	Almost always listens to, shares with, and supports the efforts of others.
3. Generally was cooperative in group activities	Is always publicly critical of the project or the work of other members of the group. Has a negative attitude towards every aspect.	Is often publicly critical of the project or the work of other members of the group. Is often negative about the task(s).	Is occasionally publicly critical of the project or the work of other members of the group. Usually has a positive attitude about the task(s).	Is rarely publicly critical of the project or the work of others. Often has a positive attitude about the task(s).	Is never publicly critical of the project or the work of others. Always has a positive attitude about the task(s).
4. Contributed to good problem-solving skills	Pretends to solve problems; Causes disruption to others work.	Does not try to solve problems or help others solve problems. Lets others do the work.	Does not suggest or refine solutions, but is willing to try out solutions suggested by others.	Refines solutions suggested by others.	Actively looks for and suggests solutions to problems.
Scoring Rubric: The Task					
5. Contributed useful ideas	Never provide idea. Free rider.	Rarely provide useful idea. Frequently provides irrelevant ideas, comments and questions.	Sometimes took initiative in generating ideas. Sometimes suggest irrelevant ideas, comments and questions.	Usually took initiative in generating ideas. Suggest ideas, comments and questions.	Always took initiative in generating creative ideas. Always suggest concrete ideas, comments and questions. Concise and clear expression of ideas.
6. Demonstrated good interest to task given	Does not complete any assigned tasks and uses others to complete his/her work.	Completed most of the individual tasks but did not assist other group members during the project.	Completed individual task and assisted other group members some times during the project.	Completed most of the assigned tasks. Volunteered to assist group members in finishing the tasks.	Completed all assigned tasks. Always assisted other group members in finishing off the tasks.
7. Prepared drafts of report in good quality	Demonstrates no focus on the topic. No organization	Demonstrates insufficient focus on the topic and provides few details. Little evidence of organization with poor transitions.	Maintains focus on the topic and provides adequate if minimal details. Logical organization with few lapses and acceptable transitions.	Maintains good focus on the topic and provides sufficient details. Logical organization that displays completeness with few lapses in transitions	Maintains exceptional focus on the topic and provides ample supporting details. Exceptional organization and provides effective transitions

References

[1] R.E. Slavin, Cooperative Learning and Cognition in Education, Elsevier Academic Press, 2011, pp. 160–166.

[2] J.E. Mills, D.F. Treagust, Engineering education – is problem-based or project-based learning the answer? Australas. J. Eng. Educ. 3 (2003) 2–16.

[3] L. Springer, M.E. Stanne, S.S. Donovan, Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: a meta-analysis, Rev. Educ. Res. 69 (1999) 21–51.

[4] P.C. Wankat, R.M. Felder, K.A. Smith, F.S. Oreovicz, The scholarship of teaching and learning in engineering, in: Disciplinary Styles in the Scholarship of Teaching and Learning: Exploring Common Ground, 2002, pp. 217–237.

[5] K.P. Nepal, An approach to assign individual marks from a team mark: the case of Australian grading system at universities, Assess. Eval. High. Educ. 37 (2012) 555–562.

[6] R. Conway, D. Kember, A. Sivan, M. Wu, Peer assessment of an individual's contribution to a group project, Assess. Eval. High. Educ 18 (1993) 45–56.

[7] L.K.Y. Lim, Some refinements on peer assessment of group projects, Assess. Eval. High. Educ 26 (2001) 5–18.

[8] D.J. Boud, W.H. Holmes, Self and peer marking in an undergraduate engineering course, IEEE Trans. Educ. 24 (1981) 267–274.

[9] T.Y. Chang, Y.T. Chen, Cooperative learning in E-learning: a peer assessment of student-centered using consistent fuzzy preference, Expert Syst. Appl. 36 (2009) 8342–8349.

[10] N. Falchikov, Peer feedback marking – developing peer assessment, Innov. Educ. Train. Int. 32 (1995) 175–187.

[11] F. Lopez-Real, Y.P.R. Chan, Peer assessment of a group project in a primary mathematics education course, Assess. Eval. High. Educ. 24 (1999) 67–79.

[12] W. Cheng, M. Warren, Making a difference: using peers to assess individual students' contributions to a group project, Teach. High. Educ. 5 (2000) 243–255.

[13] C.C. Chang, K.H. Tseng, P.N. Chou, Y.H. Chen, Reliability and validity of Web-based portfolio peer assessment: a case study for a senior high school's students taking computer course, Comput. Educ. 57 (2011) 1306–1316.

[14] K.J. Topping, Peer assessment, Theory Pract. 48 (2009) 20–27.

[15] K.J. Topping, Peers as a Source of Formative and Summative Assessment Sage Handbook of Research on Classroom Assessment, 395–412, 2013.

[16] C. Garrison, M. Ehringhaus, Formative and Summative Assessments in the Classroom, 2007 http://schools.nyc.gov/NR/rdonlyres/33148188-6FB5-4593-A8DF-8EAB8CA002AA/0/2010_11_Formative_Summative_Assessment.pdf

[17] J.B. Yang, Y.M. Wang, D.L. Xu, K.S. Chin, The evidential reasoning approach for MADA under both probabilistic and fuzzy uncertainties, Eur. J. Oper. Res. 171 (2006) 309–343.
[18] R. Biswas, An application of fuzzy sets in students' evaluation, Fuzzy Sets Syst. 74 (1995) 187–194.
[19] J. Ma, D. Zhou, Fuzzy set approach to the assessment of student-centered learning, IEEE Trans. Educ. 43 (2000) 237–241.
[20] R.C.W. Kwok, D. Zhou, Q. Zhang, J. Ma, A fuzzy multi-criteria decision making model for IS student group project assessment, Group Decis. Negotiat. 16 (2007) 25–42.
[21] S.M. Bai, S.M. Chen, Evaluating students' learning achievement using fuzzy membership functions and fuzzy rules, Expert Syst. Appl. 34 (2008) 399–410.
[22] I. Saleh, S.I. Kim, A fuzzy system for evaluating students' learning achievement, Expert Syst. Appl. 36 (2009) 6236–6243.
[23] K.M. Tay, C.P. Lim, A fuzzy inference system-based criterion-referenced assessment model, Expert Syst. Appl. 38 (2011) 11129–11136.
[24] L.A. Zadeh, Fuzzy logic = computing with words, IEEE Trans. Fuzzy Syst. 4 (1996) 103–111.
[25] J.M. Mendel, D. Wu, Perceptual Computing: Aiding People in Making Subjective Judgments, Wiley–IEEE Press, 2010.
[26] D. Wu, J.M. Mendel, Aggregation using the linguistic weighted average and interval type-2 fuzzy sets, IEEE Trans. Fuzzy Syst. 15 (2007) 1145–1161.
[27] D. Wu, J.M. Mendel, Computing with words for hierarchical decision making applied to evaluating a weapon system, IEEE Trans. Fuzzy Syst. 18 (2010) 441–460.
[28] S. Han, J.M. Mendel, A new method for managing the uncertainties in evaluating multi-person multi-criteria location choices, using a perceptual computer, Ann. Oper. Res. 195 (2012) 277–309.
[29] H. Acosta, D. Wu, B.M. Forrest, Fuzzy experts on recreational vessels, a risk modelling approach for marine invasions, Ecol. Model. 221 (2010) 850–863.
[30] K.C. Chai, K.M. Tay, C.P. Lim, A new fuzzy ranking method using fuzzy preference relations, in: IEEE International Conference on Fuzzy Systems (FUZZ), 2014, pp. 1–5.
[31] A.M. Nejab, M. Mashinchi, Ranking fuzzy numbers based on the areas on the left and the right sides of fuzzy number, Comput. Math. Appl. 61 (2011) 431–442.
[32] S.M. Chen, L.W. Lee, Fuzzy multiple criteria hierarchical group decision-making based on interval type-2 fuzzy sets, IEEE Trans. Syst. Man Cybern. A 40 (2010) 1120–1128.
[33] J.M. Mendel, R.I. John, F.L. Liu, Interval fuzzy-2 fuzzy logic system made simple, IEEE Trans. Fuzzy Syst. 14 (2006) 808–821.
[34] X. Wang, E.E. Kerre, Reasonable properties for the ordering of fuzzy quantities (I) & (II), Fuzzy Sets Syst. 118 (2001) 375–405.
[35] S. Massanet, J.V. Riera, J. Torrens, E. Herrera-Viedma, A new linguistic computational model based on discrete fuzzy numbers for computing with words, Inf. Sci. 258 (2014) 277–290.
[36] S. Alonso, I.J. Pérez, F.J. Cabrerizo, E. Herrera-Viedma, A linguistic consensus model for web 2.0 communities, Appl. Soft Comput. 13 (2013) 149–157.
[37] Y.M. Wang, K.S. Chin, G.K.K. Poon, J.B. Yang, Risk evaluation in failure mode effects analysis using fuzzy weighted geometric mean, Expert Syst. Appl. 36 (2009) 1195–1207.
[38] J.J. Wang, Y.Y. Jing, C.F. Zhang, J.H. Zhao, Review on multi-criteria decision analysis aid in sustainable energy decision making, Renew. Sustain. Energy Rev. 13 (2009) 2263–2278.