A Revised Relative Approach and Empirical Study for Modeling and Predicting Teaching Effectiveness

Kun Tian
CS Department
California State University Chico
Chico, CA, USA
ktian@csuchico.edu

Jiayi Li
School of Comp. and Info. Eng.
Tianjin Normal University
Tianjin, China
3044581379@qq.com

Jiaqi Liu
School of Comp. and Info. Eng.
Tianjin Normal University
Tianjin, China
ljq02811@qq.com

Mingyuan Chen
College of Software
Tianjin Normal University
Tianjin, China
cmy990830@qq.com

Ying He
School of Management
Tianjin Normal University
Tianjin, China
heying@tjnu.edu.cn

Wen Liu
Department of Mathematics
Lamar University
Beaumont TX, USA
wliu@lamar.edu

Guiqin He
College of Software
Tianjin Normal University
Tianjin, China
hgqwork@qq.com

Huwei Liu
School of Comp. and Info. Eng.
Tianjin Normal University
Tianjin, China
sse_lhw@163.com

***Abstract*—In this paper, two revisions are proposed for the *T-Matrices* modeling framework, which provides an objective and accurate method for teaching effectiveness modeling, assessment, and prediction from a unique perspective on teaching effectiveness that captures the interactive nature of teaching and learning. The revisions have been motivated by addressing the limitations of the original framework in teaching style (mode) modeling and teaching effectiveness comparability. The revised framework is then proposed as a new version, *T-Matrices Rev*. In addition, a full scale empirical study is given to illustrate the effectiveness of *T-Matrices Rev* on teaching effectiveness prediction. The study, based on the data collected during a time span of 4 years with 7 semesters, showed promising results, which are discussed along with several insights on how to extend the framework-based approach in our future work for further developments and better applicability.**

Keywords—Educational Evaluation Method, Relative Teaching Effectiveness, Teaching Effectiveness Assessment

I. INTRODUCTION

Educational assessment [1,2,3,4,5,6,7] is a systematic process of documenting and analyzing empirical data to help students learn better. It can be either a direct assessment using data or an indirect one using references on data. The current major and common practice for teaching effective evaluation is course evaluation, a direct teaching quality assessment by students. Evaluation criteria given on a course evaluation typically include fairness in grading, communication skills, enthusiasm, flexibility, presentation skills, student engagement, etc. [8,9] Such summative assessment on teaching effectiveness has been widely criticized as not being a fair and accurate measurement [10,11,12,13,14,15,16].

Initial efforts on using a mathematics-specific framework for evaluation were found useful but ineffective at handling concepts that are not mathematics-specific [17]. A forum for synergic collaboration on researching teaching quality was proposed, as there is a lack of a common language delineating teaching effectiveness [18]. Finally, reflections on teaching quality also propelled researchers to rethink its definition, and with an empirical study, it was shown that teaching quality also depends on the composition of the student body [19].

T-Matrices modeling framework [20] has been proposed based on the relative teaching effectiveness idea [19] for teaching effectiveness modeling, assessment, and prediction. However, there are currently two applicability issues in the original framework in teaching style modeling and teaching effectiveness comparability. Addressing the issues by proposing two related revisions, we transform the original framework into its new version *T-Matrices Rev*.

An empirical study on a much fuller scale than the original one presented in [20] is also given to illustrate the effectiveness of *T-Matrices Rev* on teaching effectiveness prediction. The case study, based on the data collected during a time span of 4 years with 7 semesters, showed promising results, which are discussed along with several insights on how to extend the framework-based approach in our future work for further developments and better applicability.

The rest of the paper is organized as follows. Section 2 provides a theoretical background for *T-Matrices*. Section 3 discusses the two limitations of *T-Matrices* and proposed its new version *T-Matrices Rev* by introducing two revisions addressing the limitations. A full scale case study validating the proposed approach is given in Section 4. Discussions are given in Section 5, which concludes this paper as well.

II. BACKGROUND

In *T-Matrices*, both a student and a course are considered as a collection of elements. As for a course, its elements represent the collection of its learning objectives.

$$C = (o^1, o^2, \dots, o^n), o^i \text{ is a learning objective of } C. \quad (1)$$

For a student in general, it is a collection of properties.

$$S = (p^1, p^2, \dots, p^l),\ p^i \text{ is a property of } S. \quad (2)$$

For a student taking a course, its elements are defined as properties that are of relevance to the course. Different courses can reference different properties of a student.

$$S_C = (p^1, p^2, \dots, p^m),\ p^i \textit{ is a property of } S \textit{ and is relevant to } C. \quad (3)$$

T-Matrices represents them as vectors.

$$S_c{}^T = \begin{bmatrix} p^1 \\ p^2 \\ \cdot \\ \cdot \\ \cdot \\ p^m \end{bmatrix}, \quad C^T = \begin{bmatrix} o^1 \\ o^2 \\ \cdot \\ \cdot \\ \cdot \\ o^n \end{bmatrix}, \quad S^T = \begin{bmatrix} p^1 \\ p^2 \\ \cdot \\ \cdot \\ \cdot \\ p^l \end{bmatrix} \quad (4)$$

When a teacher T teaches a course C, his/her teaching is defined as a matrix.

$$E_{TC} = \begin{bmatrix} E_{TC}^{11} & E_{TC}^{21} & . & E_{TC}^{n1} \\ E_{TC}^{12} & E_{TC}^{22} & . & E_{TC}^{n2} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ E_{TC}^{1m} & E_{TC}^{2m} & . & E_{TC}^{nm} \end{bmatrix}, \ \textit{where} \begin{cases} n = |C| \\ m = |S_C| \end{cases} \quad (5)$$

In the above matrix, n is the number of learning objectives in C and m is the number of properties in SC, and each of ETC's components is a function that takes a property of SC as the only parameter. Teaching a course C to a student S by a teacher T is considered as a process that applies T's teaching to S's relevant properties, and the process generates results in terms of the degree of completeness of the course's various learning objectives.

$$S_c \circ E_{TC} = \begin{bmatrix} E_{TC}^{11}(p^1) + & E_{TC}^{12}(p^2) + & . & E_{TC}^{1m}(p^m) \\ E_{TC}^{21}(p^1) + & E_{TC}^{22}(p^2) + & . & E_{TC}^{2m}(p^m) \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ E_{TC}^{n1}(p^1) + & E_{TC}^{n2}(p^2) + & . & E_{TC}^{nm}(p^m) \end{bmatrix}^T = R_{TC}^S = \begin{bmatrix} q^1 \\ q^2 \\ \cdot \\ \cdot \\ \cdot \\ q^n \end{bmatrix}^T \quad (6)$$

$$\textit{where } q^i \in [0,1] \quad (7)$$

By the above formula, T-Matrices recognizes that the completeness of a learning objective is the collective consequences of a teacher's various efforts working on a student's relevant properties toward the objective. It also recognizes a course's teaching results as the collection of the degree of completeness of all its learning objectives (a.k.a. learning outcomes) by the teacher and his/her student(s). (Discussions on the conformities of the models to reality are given in the last section.)

Teaching effectiveness (quality) is defined as the L2 norm of the results, as shown in below.

$$Q_T^C(S_c) = \| R_{TC}^S \|_2 \quad (8)$$

For a group of students **S**, the collective teaching result is defined as follows. One thing to note is that the collective result is in a matrix form.

$$\boldsymbol{S_c} = \begin{bmatrix} S_c{}^1 \\ S_c{}^2 \\ \cdot \\ \cdot \\ \cdot \\ S_c{}^k \end{bmatrix} \qquad \boldsymbol{R}_{TC}^{\boldsymbol{S}} = \boldsymbol{S_c} \circ E_{TC} \quad (9)$$

Subsequently, the collective teaching effectiveness is defined as the Frobenius norm of the results.

$$Q_T^C(\boldsymbol{S_c}) = \| \boldsymbol{R}_{TC}^{\boldsymbol{S}} \|_F \quad (10)$$

By the above formula, T-Matrices recognizes that the completeness of a learning objective is the collective consequences of a teacher's various efforts working on a student's relevant properties toward the objective. It also recognizes a course's teaching results as the collection of the degree of completeness of all its learning objectives (a.k.a. learning outcomes) by the teacher and his/her student(s). (Discussions on the conformities of the models to reality are given in the last section.)

Teaching effectiveness (quality) is defined as the L2 norm of the results, as shown in below.

$$Q_T^C(S_c) = \| R_{TC}^S \|_2 \quad (11)$$

For a group of students **S**, the collective teaching result is defined as follows. One thing to note is that the collective result is in a matrix form.

$$\boldsymbol{S_c} = \begin{bmatrix} S_c{}^1 \\ S_c{}^2 \\ \cdot \\ \cdot \\ \cdot \\ S_c{}^k \end{bmatrix} \qquad \boldsymbol{R}_{TC}^{\boldsymbol{S}} = \boldsymbol{S_c} \circ E_{TC} \quad (12)$$

Subsequently, the collective teaching effectiveness is defined as the Frobenius norm of the results.

$$Q_T^C(\boldsymbol{S_c}) = \| \boldsymbol{R}_{TC}^{\boldsymbol{S}} \|_F \quad (13)$$

The details on data collection, normalization, and formatting for *T-Matrices* is given in [20].

III. T-MATRICES REV

A. T-Matrices Problems

T-Matrices has the following problems.

First, though T-Matrices recognizes the fact that many teachers are able to season their ways of teaching a course, which means they usually have more than one teaching mode for a course, in notation it doesn't differentiate between them and hence may bring ambiguities.

Second, the definition of teaching effectiveness is not normalized in T-Matrices. It means that courses with more learning objectives tend to result in bigger values on teaching results and effectiveness. It also makes the values of teaching effectiveness not comparable for the different sections (of a same course) that have different numbers of students.

B. T-Matrices Revisions: T-Matrices Rev

To address the first problem, we provide an extension to the teaching mode definition in T-Matrices.

$$E_{TC}{}^{\alpha}, where\ \alpha\ denotes\ T's\ teaching\ mode\ index$$

With the above new notation, assuming that the models are instantiated (probably as a result of training) and the data on students' properties are available, we can use them to predict teaching effectiveness, which brings various strategic benefits. An example is given as follows.

When preparing for a course for a target set of students, a teacher should select his/her teaching mode that will generate the best teaching effectiveness.

$$\hat{\alpha} = \underset{E_{TC}{}^{\alpha}}{arg\,max}\ Q_T^C(\boldsymbol{S_c}) \tag{14}$$

In addition, we can find a course schedule that maximize the grand teaching effectiveness of all the courses given in a semester similarly as follows.

$$\hat{P} = \underset{P}{arg\,max}\ \sum_C \theta_C Q_{P(C)}^C(\boldsymbol{S}(C)_{\boldsymbol{C}}) \tag{15}$$

$$where \begin{cases} P \in \begin{Bmatrix} assignment \\ permutations \end{Bmatrix} \\ P(C){:}\ C's\ instroctor\ by\ P \\ \theta_C{:}\ C's\ weighting\ factor \\ \boldsymbol{S}(C){:}\ students\ in\ C \end{cases} \tag{16}$$

To address the second problem, we propose to introduce a coefficient vector for C's various learning outcomes. The vector stands for the weights of the learning outcomes.

$$C^T = \begin{bmatrix} o^1 \\ o^2 \\ \cdot \\ \cdot \\ \cdot \\ o^n \end{bmatrix},\ CO^T = \begin{bmatrix} co^1 \\ co^2 \\ \cdot \\ \cdot \\ \cdot \\ co^n \end{bmatrix},\ where\ \sum_i co^i\ = 1 \tag{17}$$

Consequently, the calculation on teaching effectiveness is revised.

$$Q_T^C(S_c) = R_{TC}^S \cdot CO^T \tag{18}$$

Teaching effectiveness for a class of students is revised as follows.

$$Q_T^C(\boldsymbol{S_c}) = \frac{\sum_S Q_T^C(S_c)}{|\boldsymbol{S_c}|},\quad where\ S_c \in \boldsymbol{S_c} \tag{19}$$

Using the above definitions, the values of teaching effectiveness for different sections of the same course (or for the same course offered in different semesters) are normalized into the same range. Therefore, they are comparable with each other.

IV. EMPRICAL STUDY

This section presents a case study, which also shows that our approach can be used to identify an instructor's new teaching mode of a course.

A. Target Course Settings

We analyzed the academic records of 12 Software Engineering course sections from 7 semesters from 2017 to 2021, which involves 588 students in the 13 sections taught by a same instructor at Tianjin Normal University (TJNU). Using the T-Matrices models defined in Section 3, we instantiated the models as follows.

The outcomes for the course has been established previously at TJNU and are given in Table 1, which consults the Software Engineering Body of Knowledge (SWEBOK) [21]. In addition, for practical reasons and objectivity, we defined the student properties referenced by this course using the students' performance scores on all the relevant prerequisite courses. The only exception is p8, which came from subjective assessments by the administrative stuff termed as class monitors, who, in Chinese universities, typically monitor and manage students closely on campus. Each class has its own monitor, who is very knowledgeable about and familiar with his/her own students.

TABLE I. THE LEARNING OBJCTIVES OF TJNU SOFTWARE ENGINEERING COURSE & REFERENCED STUDENT PROPERTIES

SE Learning Objectives	
ID	***Categories***
o^1	Software Crisis
o^2	Software Definition
o^3	Software Life Cycle Models
o^4	Requirements Analysis
o^5	Object-Oriented Programming
o^6	Software Architecture Design
o^7	Software Detailed Design
o^8	UML
o^9	Software Construction
o^{10}	Software Testing
o^{11}	Software Maintenance
o^{12}	Software Project Management
o^{13}	Software Project Documentation
o^{14}	Software Configuration Management
o^{15}	Software Engineering Code Of Ethics
o^{16}	Project Based Software Development Experiences

SE Referenced Student Properties		
ID	***Major Categories***	***Minor Categories***
p^1	Mathematical and Analytical Capability	College Algebra I
p^2		College Algebra II
p^3		Discrete Mathematics
p^4	Computer Fundamentals	Operating System
p^5		Computer Organization and Architecture
p^6	Programming Experiences	C++ Programing Or Java Programming
p^7	Technical Writing and the Abilities to Express and Understand	Natinoal College Entrance Examination on Chinese Literature and Language
p^8	Capability of Teamwork	Class Monitor's Evaluations in their Profile Archives

In order to normalize students' performance data on gradable items, we retrieved the weights of the learning objectives, which also has been historically established and remained effective for this course at TJNU, as shown in Table 2. An example of final grade calculation is given in Table 3.

TABLE II. WEIGHTS DISTRIBUTIN OF THE SOFTWARE ENGINEERING OBJECTIVES AT TJNU

ID	*Categories*	*Pcnt.*
o^1	Software Crisis	2
o^2	Software Definition	2
o^3	Software Life Cycle Models	10
o^4	Requirements Analysis	10
o^5	Object-Oriented Programming	15
o^6	Software Architecture Design	2
o^7	Software Detailed Design	2
o^8	UML	15
o^9	Software Construction	10
o^{10}	Software Testing	2
o^{11}	Software Maintenance	2
o^{12}	Software Project Management	2
o^{13}	Software Project Documentation	2
o^{14}	Software Configuration Management	2
o^{15}	Software Engineering Code Of Ethics	2
o^{16}	Project Based Development Experiences	20

TABLE III. FINAL GRADE CALCULATION EXAMPLE

ID	*Scores*	*Pcnt.*	*Weighted Scores*
o^1	100	2	2
o^2	100	2	2
o^3	80	10	8
o^4	100	10	10
o^5	70	15	10.5
o^6	60	2	1.2
o^7	90	2	1.8
o^8	80	15	12
o^9	75	10	7.5
o^{10}	100	2	2
o^{11}	100	2	2
o^{12}	50	2	1
o^{13}	50	2	1
o^{14}	60	2	1.2
o^{15}	100	2	2
o^{16}	85	20	17
Cumulative Grade			81

Shown in Table 4 is a percentile to letter grade conversion at TJNU. However, there is a special case for p7, as students from Jiangsu Province of China were assessed using a special points system on National College Entrance Examination on Chinese Literature and Language. It is a 120 pts based one rather than 150 pts based one, which is used by all the other parts of China. The corresponding special grade conversion is also provided in Table 5. The reason for letter grade conversion is because that we argue that we should emphasize the qualitative rather than quantitative significance in grade, e.g. the grades of 86 and 88 both indicate a performance level that is “rather good” even if they are numerically different.

TABLE IV. TJNU OFFICIAL PERCENTILE TO LETTER GRADE CONVERSIONS

Prct.	100-93	92-90	89-87	86-84	83-80
Letr.	A	A-	B+	B	B-
Pcnt.	79-75	74-70	69-66	65-60	59-0
Letr.	C+	C	C-	D	F

TABLE V. LETTER GRADE CONVERSIONS FOR P7

For Students Not From Jiangsu				
150-140	139-135	134-131	130-126	125-120
A	A-	B+	B	B-
119-113	112-106	105-99	98-90	89-0
C+	C	C-	D	F
For Students From Jiangsu				
120-112	111-108	107-105	104-101	100-96
A	A-	B+	B	B-
95-90	89-84	83-80	79-72	71-0
C+	C	C-	D	F

B. *Data Training and Verification Results*

We first used the data of 7 sections of the course (317 students in total) between 2017 and 2019, of which 5 sections were used for training the teaching simulation engine for the instructor. The last two sections (Set 1 and 2) were used for testing the engine. Afterwards, we used the data of 6 sections of course (271 students in total) between 2020 and 2021 for testing it (Set 3 to 8). One thing to note is that the last 6 sections of the course were taught online due to the impacts of Covid 19, even if the final exams were structured the same as before it and were given on site. For training the data, we used two-layer feed-forward neural networks and the classical BP algorithm (learning rate $\eta = 0.2$).

TABLE VI. PREDICTION RESULTS BY THE SIMULATING ENGINE (SET 1 & 2, F19)

Diff.	P_{Crct}	P_{Half}	P_{One}	P_{Err}	*Samples*
	NA	½ Ltr	1 Ltr	> 1 Ltr	
Set 1	32	5	1	3	41
Set 2	34	6	3	3	46
Sum	66	11	4	6	87
Pcnt.	75.86	12.64	5.75	6.90	100

The predication results for Set 1 and 2 are given in Table 6. In addition, we defined the credibility of the grade prediction as follows. Because we argue that certain amount of error should be allowed, as the predication of grade range is more meaningful and useful than the exact grade.

$$Credibility\ of\ Prediction\ = P_{Crct} + P_{Half} \tag{20}$$

From Table 6, the credibility of prediction for Set 1 and 2 (F19 semester) is 88.50%. The prediction results for the course in the other three semesters (S20, F20, and S21) are given in Table 7, 8, and 9. For S20, the credibility of prediction is 63,64%. For F20, it is 57.05%. For S21, it is 68.89%. As shown, there are noticeable drops in the predication credibility numbers.

TABLE VII. PREDICTION RESULTS BY THE SIMULATING ENGINE (SET 3, 4, & 5, S20)

Diff.	P_{Crct}	P_{Half}	P_{One}	P_{Err}	Samples
	NA	½ Ltr	1 Ltr	> 1 Ltr	
Set 3	15	18	2	8	46
Set 4	13	9	11	11	44
Set 5	19	10	8	10	42
Sum	47	37	21	29	132
Pcnt..	35.61	28.03	15.91	21.97	100

TABLE VIII. PREDICTION RESULTS BY THE SIMULATING ENGINE (SET 6, F20)

Diff.	P_{Crct}	P_{Half}	P_{One}	P_{Err}	Samples
	NA	½ Ltr	1 Ltr	> 1 Ltr	
Set 6	16	12	10	11	49
Sum	16	12	10	11	49
Pcnt.	32.65	24.49	20.41	22.45	100

TABLE IX. PREDICTION RESULTS BY THE SIMULATING ENGINE (SET 7 & 8, S21)

Diff.	P_{Crct}	P_{Half}	P_{One}	P_{Err}	Samples
	NA	½ Ltr	1 Ltr	> 1 Ltr	
Set 7	15	18	2	8	43
Set 8	19	10	8	10	47
Sum	34	28	10	18	90
Pcnt.	37.78	31.11	11.11	20	100

We believed that the drops were caused by the change of teaching mode by the professor. To test it, we used the same training method to train another teaching effectiveness prediction engine using Set 3 to 6 and tested the new engine using Set 7 and 8. The results are shown in Table 10.

TABLE X. PREDICTION RESULTS BY THE NEW SIMULATING ENGINE (SET 7 & 8, S21)

Diff.	P_{Crct}	P_{Half}	P_{One}	P_{Err}	Samples
	NA	½ Ltr	1 Ltr	> 1 Ltr	
Set 7	30	6	2	5	43
Set 8	37	4	4	2	47
Sum	67	10	6	7	90
Pcnt.	74.44	11.11	6.67	7.78	100

As shown in Table 10, the credibility of the new prediction is 85.55%. It corroborates, at least partially, that another teaching mode than that for Set 1 and 2 (F19) students have been applied to Set 3 to 9 students, and the new mode has been captured and successfully reproduced by data training using information from the affected sections. Moreover, this case also shows that our revisions done to T-Matrices on separate modeling of teaching mode is necessary to identify changes in teaching (course delivery). Also as shown in this case study, if the prediction results' credibility is significantly low (lower than 69% in this case) and there is no apparent change in the composition of students' body nor for the instructor assigned, it is very likely that there has been a major change in the delivery of the course (change of teaching mode) as a result of various possible changes in course settings (e.g. onsite vs. online). As a result, a new teaching simulation engine should be built to capture the new mode for better prediction results. Finally, the new predictions achieved the credibility value bigger than 85%, which shows that our approach is, to a certain extent, promising. Better values may be obtained by tuning the data training process using better AI techniques.

V. DISSCUSIONS AND CONCLUSIONS

For the matrix-based teaching mode model in T-Matrices Rev, computation toward the degree of completeness of a learning objective is based on a linear composition of the results by its various sub-functions, each of which takes only one student-property as a parameter. It partially implies that these properties are independent. Considering the possible interferences among student-properties, other computation methods can also be applied. However, we argue that the teaching mode model should be considered mainly as a transformation that maps student-properties to the degrees of completeness of a course's various learning objectives. In the absence of exact methods in clear mathematical forms that compute the results, it can be used as an avatar for them, which will be derived into a teaching mode simulation engine. In this sense, the teaching mode model can be considered as a mapping from student properties to a set of values that correspond to the completeness of each of the course's learning outcomes.

$$E_{TC}{}^{\alpha}(S_C) = Completeness_C(cpt_1, cpt_2, \dots, cpt_n) \quad (21)$$
$$where \begin{cases} |C| = n \\ 0 \le cpt_i \le 1 \end{cases}$$

As teaching is an interactive process, it also may affect student properties. In other words, students' properties may be different after taking a course. In this sense, a teaching mode can be extended so that it produces a set of updated student properties.

$$E_{TC}{}^{\alpha}(S_C) = \left(S_C{}', Completeness_C(cpt_1, cpt_2, \dots, cpt_n)\right) \quad (22)$$

Using a student profile as a collection of properties (that may be stored on a medium like blockchain), we can predict student-properties' changes in a way that is similar to teaching effectiveness prediction, and it is shown in Figure 1.

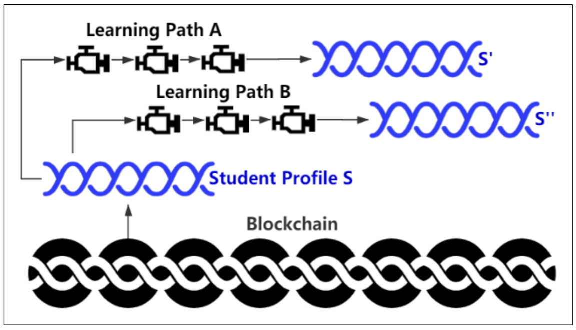

Figure. 1. The Illustration on Predictions of Alternative Student Profiles.

Computations using different learning paths, which respectively consist of a series of teaching modes, generate different sets of resulting properties. We can decide which path is the most desirable based on possible preferences.

Determining student-properties and establishing the correspondences between student-properties to course requirements present challenges. It also represents the "experimental science" aspect of this research. We suggest establishing a forum and/or a platform for educational practitioners, experts, and related professionals to collaborate on the experiments synergically. We also suggest using a bottom-up approach, in which the determinations and establishments concerning individual courses take sprecedence. As our knowledge of them deepens, the results can be generalized, combined, and/or synthesized to build (or at least to elicit insights on) those on a higher level, e.g., program level, discipline level, institutional level, etc. We should go on this academic expedition in the future.

Using *T-Matrices Rev* definition on teaching effectiveness, it appears that the quality of teaching is measured objectively by "how well students complete learning outcomes". This definition can be extended to cover more angels of views. For example, if we want the definition to cover other non-objective criteria such as "student satisfaction", a straightforward extension would be adding "student satisfaction" to the course's learning outcomes. If doing so is deemed as against the objective/academic principle of categorizing a course's learning outcomes, we can make another model (vector) that include all of these subjective criteria (as "outcomes") and use it the same way we do with a course's model (vector) of learning outcomes to compute the quality of teaching in such criteria.

To summarize, we have proposed a semi-formal approach to model, compute, and predict teaching effectiveness in a way that captures the interactive nature of teaching/learning. The approach can be used to differentiate an instructor's various teaching modes of a course. A case study using the approach was given. The potential problems are discussed, which we would address in the future.

REFERENCES

[1] Mary J. Allen, Assessing Academic Programs in Higher Education. San Francisco: Josei-Bass, 2003.

[2] Susan M. Brookhart, Anthony J. Nitko, Educational Assessment of Students (8th Ed), Perason, 2018.

[3] Robert Nelson, Phillip Dawson, A Contribution to the History of Assessment: How a Conversation Simulator Redeems Socratic Method. *Assessment & Evaluation in Higher Education*, vol. 29, issue 2, pp. 195-204.

[4] Linda Suskie, Assessing Student Learning: A Common Sense Guide (3rd Ed). MA: Anker, 2018.

[5] Kathryn Parker Boudett, Elizabeth A. City, Richard J. Murnane, Data Wise, Revised and Expanded Edition: A Step-by-Step Guide to Using Assessment Results to Improve Teaching and Learning Revised and Expanded Edition. Harvard Education Press, 2013.

[6] Michael Strong, The Highly Qualified Teacher. NY: Teachers College Press, 2011.

[7] Linda Darling-Hammond, Peter Youngs, Defining "Highly Qualified Teachers." What Does Scientifically-Based Research Actually Tell Us? *Educational Researcher*, vol. 31, issue 9, pp. 13-25, 2012.

[8] Rahman Kaleel, Learning from Your Business Lectures: Using Stepwise Regression to Understand Course Evaluation Data. *Journal of American Academy of Business*, vo. 19, issue 2, pp. 272-279, 2006.

[9] Chong Kim, Earl Damewood, Norma Hodge, Professor Attitude: Its Effect on Teaching Evaluations. *Journal of Management Education*, vol. 24, issue 4, pp. 458–473, 2000.

[10] J. Scott. Armstrong, *Natural Learning in Higher Education*, Retrieved from https://faculty.wharton.upenn.edu/wp-content/uploads/2014/09/Natural-Learning-in-Higher-Education_2.pdf. Accessed: Feb 5, 2021.

[11] M. E. Schaff, B. R. Siebring, A Survey of the Literature Concerning Student Evaluation of Teaching, *Journal of Chemical Education*, vol. 51, issue 3, pp. 150-151, 1974.

[12] Michael A. McPherson, R. Todd Jewell, Myungsup Kim, What Determines Student Evaluation Scores? A Random Effects Analysis of Undergraduate Economics Classes. *Eastern Economic Journal*, vol. 35, issue 1, pp. 37–51, 2009.

[13] Laura Langbein, Management by Results: Student Evaluation of Faculty Teaching and the Mis-Measurement of Performance. *Economics of Education Review*, vol. 27, issue 4, pp. 417–428, 2006.

[14] Scott E. Carrell, James E. West, Does Professor Quality Matter? Evidence from Random Assignment of Students to Professors. *Journal of Political Economy*, vol. 118, issue 3, pp. 409–432, 2010.

[15] Marks Pia, Silent Partners: Student Course Evaluations and the Construction of Pedagogical Worlds. *Discourse and Writing/Rédactologie*, vol. 24, pp. 1-32, 2012

[16] Linda J. Grahama, Sonia L.J. Whitea, Kathy Cologonb, Robert C. Piantac, Do Teachers' Years of Experience Make a Difference in the Quality of Teaching? *Teaching and Teacher Evaluation*, vol. 96, Article Number 103190, 2020.

[17] Litkea Erica, Boston Melissa, Walkowiak Temple A. (2021). Affordances and Constraints of Mathematics-Specific Observation Frameworks and General Elements of Teaching Quality. Studies in Educational Evaluation, 68, Article 100956. https://doi.org/10.1016/j.stueduc.2020.100956.

[18] Charalambos Y. Charalambous, Anna-Katharina Praetorius, Creating a Forum for Researching Teaching and its Quality More Synergistically. *Studies in Educational Evaluation*, vol. 67, Article Number 100894, 2020.

[19] Benjamin Fauth, Cansu Atlay, Hanna Dumont, Jasmin Decristan. Does What You Get Depend on Who You Are with? Effects of Student Composition on Teaching Quality. *Learning and Instruction*, vol. 71, Article Numer 101355, 2021.

[20] Kun Tian, Wen Liu, Ying He, Ming Yang, Danhua Zhao, On the Modeling and Prediction of Teaching Effectiveness with Machine Learning. Paper Presentation at the 15th Internation Conference on Computer Science and Education, Delft, Netherlands, 2020.

[21] Bourque P. and Fairley R.E. eds., Guide to the Software Engineering Body of Knowledge, Version 3.0. IEEE Computer Society, 2014; www.swebok.org.