A Study of Multiple Teacher Evaluation in the United States Based on Artificial Intelligence: Comparison of Danielson and Marzano Evaluation Models

Di Yuan[a]
Educational Studies, Fujian Normal University Fuzhou, China

Keywords: The United States, AI, Modern Information Technology, Teacher Evaluation, Danielson, Marzano, Teaching and Learning.

Abstract: Since the 1980s, the research on educational evaluation in the United States ushered in a "multi-model period", and corresponding teacher evaluation models have emerged to seek a symbiosis between the development of teachers' professionalism and the enhancement of students' academic achievement. This paper takes the Danielson Framework for Teaching and the Marzano Teacher Evaluation Model by modern information technology as examples, analyzing their backgrounds, model content, similarities, and differences. It aims to provide references for reflecting on and promoting current teacher evaluation practices in China.

1 INTRODUCTION

The 1980s was the fourth phase in American educational evaluation history, following the "pre-educational measurement period", "educational measurement period" and "goal analysis period", namely, the "multi-model period"' (Wang 2018). Influenced by this climate, Danielson, and Marzano, the two most widely used teacher evaluation models were later proposed in the United States. Since its inception, the Danielson Framework has been adopted by many educators in the United States and worldwide. It is widely recognized both in the field of teacher evaluation theory and in classroom practice (Wu et al., 2019). In 2009, the Measures of Effective Teaching (MET) project, funded by the Bill & Melinda Gates Foundation, used the Danielson Framework to analyze over 23,000 course videos (Wu, Wu, and Zhang, 2020). It has also been used to observe classroom teaching in the UK, Germany, South Korea, and South Africa. The Marzano Teacher Evaluation Model, which is used by school districts in over 50 states in the United States, is one of four evaluation models developed by Marzano in collaboration with the Learning Sciences International platform. This study will discuss these two models in detail, with their backgrounds, model contents, similarities, and differences.

2 THE BACKGROUND OF DANIELSON AND MARZANO TEACHER EVALUATION MODEL BASED ON ARTIFICIAL INTELLIGENCE

2.1 Background of the Danielson Framework for Teaching

The Danielson Framework for Teaching (DFT) is an assessment tool developed by Professor Charlotte Danielson's team in the United States. It is a comprehensive analysis of teachers' professional development before, during and after the class, which has been published as Enhancing Professional Practice: A Framework for Teaching.

For states and local agencies to have a common teacher certification system, the centre of Educational Testing Service (ETS) undertook a large-scale

[a] https://orcid.org/0000-0002-5399-6310

Yuan, D.
A Study of Multiple Teacher Evaluation in the United States Based on Artificial Intelligence: Comparison of Danielson and Marzano Evaluation Models.
DOI: 10.5220/0011909400003613
In *Proceedings of the 2nd International Conference on New Media Development and Modernized Education (NMDME 2022)*, pages 221-228
ISBN: 978-989-758-630-9

research project -- the Praxis Series: Praxis I. Assessment of Students Skills; Praxis II. Assessment of Subject Matter Literacy; and Praxis III. Assessment of Classroom Performance. The significance of the first two is to award primary teaching qualifications to those who pass them, while the latter revolves around the assessment of teachers' actual teaching skills and classroom performance (Zhou 2017). Professor Danielson and her team were responsible for the development of Praxis III. However, during the study, Danielson was adamant that Praxis III was not only a tool to assess classroom competence, but also to improve teachers' professional development. Subsequently, based on Praxis III, Danielson and her team completed Danielson Framework for Teaching after continuous refinement and extended it to support teachers' professional development in teaching practices in states across the United States (Charlotte, Danielson 2013).

2.2 Background of the Marzano Teacher Evaluation Model Based on Artificial Intelligence

The Marzano Teacher Evaluation Model (MTEM) was developed for three general reasons. First, since the 1980s, neoliberalism has prevailed, and the educational process has become more concerned with cost-cutting, standard-setting, and educational output (Lu 2006). Educational outputs are reflected in the increased focus on students' test scores and have eventually been applied to teacher evaluation (Hursh 2005). Secondly, the Council for the Accreditation of Educator Preparation (CAEP) proposed in late 2012 that standards for educational evaluation systems should provide multiple assessment indicators. In response to the need for diversified teacher evaluation, Marzano and his team created a new teacher evaluation model by distilling and summarizing the core competencies of teachers through scientific evidence, which provided multiple options for the development of teacher professionalization.

Thirdly, Marzano has always been passionate about research on classroom practice, teacher evaluation, and school leader assessment, and has been committed to effectively applying the latest theories to classroom practice (Larsen 2015). The Marzano Teacher Evaluation Model was developed by Marzano and his team based on years of research, with key findings such as, What Works in Schools, Classroom Instruction that Works, Classroom Management that Works, Classroom Assessment and Grading the Work, The Art and Science of Teaching, Effective Supervision: Supporting the Art and Science of Teaching (Marzano Center, 2015).

In short, the emergence of the Marzano Model is closely linked to the deepening of neoliberalism and accountability in public education in the United States, the aspirations of the American Council for the Accreditation of Teacher Education for pluralistic teacher evaluation, and Marzano's tireless efforts.

2.3 Artificial Intelligence Application Areas

2.3.1 Natural Language Processing

Natural language processing consists of two parts: natural language understanding and natural language generation. The function of natural language understanding technology is to enable computers to understand the meaning of natural language text, and natural language generation technology is to enable computers to express given ideas and intentions in natural language text. The purpose of developing natural language processing technology is to prevent people from spending a lot of time and effort to learn various obscure computer languages, and to allow people to use the natural language they are most familiar with and accustomed to in order to achieve natural language communication between humans and computers.

2.3.2 Big Data Analytics

Big data analytics is the ability to process data of all types and shapes and to capture the information value of massive and high-growth data in a new processing model. By collecting, storing and mining data, big data analytics can help human beings find the correlation between known variables and make scientific and intelligent decisions accordingly. There is a large amount of data in the process of education and teaching, and the targeted construction of AI analysis models can help teachers identify the shortcomings in teaching and provide improvement solutions by analyzing these data with the help of big data analysis technology. The level of application of artificial intelligence in education and teaching depends on the upgrading and improvement of big data analysis technology.

3 THE FRAMEWORK OF THE DANIELSON AND MARZANO TEACHER EVALUATION MODEL

3.1 A Framework of Danielson Model of Teacher Evaluation

Danielson and her team have altered the framework four times in the last 20 years. The original version, published in 1996, included 22 dimensions with detailed descriptions. The 22 elements were then separated into four domains in the second edition in 2007 and the third edition in 2011: 'planning and preparation,' 'classroom environment,' 'classroom teaching,' and 'professional duties.' The 2013 edition of the framework has undergone significant revisions, drawing on several teaching concepts from the 2010 Common Core State Standards (CCS), with a focus on student initiative, coherent instructional design, and the use of formative assessment (Robert J., 2013).

Table 1. The Danielson Teacher Evaluation Model

Dimensions	Contents
Board 1: Planning and preparation	1a: Mastery of subject content and teaching methods 1b: Getting to know students 1c: Establishing teaching objectives 1d: Understanding teaching resources 1e: Designing coherent teaching 1f: Designing a student assessment system
Board 2: Classroom environment	2a: Creating a classroom environment of mutual respect and harmony 2b: Building a learning culture 2c: Managing the teaching process of the course 2d: Managing student behaviour 2e: Organizing physical space
Board 3: Classroom teaching	3a: Communicating with students 3b: Use questioning and discussion techniques 3c: Engaging students in learning 3d: Using evaluation in teaching 3e: Flexible and responsive
Board 4: Professional duties	4a: Reflective Teaching 4b: Keeping accurate records 4c: Communicating with students 4d: Participation in professional groups' families 4e: Professional growth and development 4f: Reflects professional qualities

Taking the fourth edition as an example (see Table 1), four domains, each containing 5-6 dimensions, for a total of 22 dimensions and 76 subelements (Hunzicker, 2017). To help users better understand and use the model, each of these 22 dimensions is divided into four levels: excellent, proficient, basic, and unqualified.

The Planning and Preparation part's components outline how teachers plan pupils' learning, or how they create instruction. As shown in Table 1, this block covers six dimensions: strong subject knowledge and pedagogy, understanding of students' learning, setting scientific teaching objectives, exploitation of teaching resources, implementation of intrinsically orderly classroom activities, and use of students' formative assessment.

The classroom is the most central part of the Danielson's framework, where the planned and prepared instructional design is put into practice and the direct activities of teaching and learning take place in a meaningful way. Teachers who excel in this area are skilled at teaching, ask thought-provoking questions, respond to students in a timely manner and shine with teaching wisdom, on the other hand, students are free to express their personal views and immerse themselves in classroom learning.

The classroom environment section describes all aspects of the classroom environment that are conducive to student learning. A good classroom environment is one in which teachers and students work together to create a relaxed and respectful classroom environment; in which with good and safe classroom furniture; in which students are actively engaged in learning; in which teaching and learning flows efficiently; in which the rules for student behaviour are

clear; and in which the classroom is set up in a way that helps students to develop logical thinking.

Professional responsibility is a crucial step for teachers from novices to experts, a core attribute of teachers' professional development. New teachers are expected to begin their careers by keeping records of student performance, reflecting on their teaching, and balancing between work and home. Once they have gained experience in teaching, they can then shift their focus to peer collaboration, work on building professional communities, focus on their own professional development, and pave the way for growth into expert teachers.

It is important to note that although each section operates independently of the others, but they're also interactive. In other words, Board 1 (Planning and preparation) prepares Board 3 (Classroom teaching) for successful implementation; Board 2 (Classroom environment) creates a harmonious external environment for Board 3 (Classroom teaching); and Board 4 (Professional responsibilities) is the value of the first three Boards and the final destination for teacher professional development.

3.2 Framework of the Marzano Teacher Evaluation Model Based on AI

The Marzano Teacher Evaluation Model has undergone two stages: the first stage was the Marzano Causal Teacher Evaluation Model (CTEM) in 2014, the second stage was the Marzano Focused Teacher Evaluation Model (FTEM) in 2017. The core feature of the Causal Model is to explore the correlation between teachers' teaching strategies and students' academic achievement. It divides teachers' work into four domains and 60 elements, each of which has some strong or weak correlation with student achievement. However, it lacks of clear and streamlined standards for teacher evaluation in real life. Later, the Focused Model simplified the previous Causal Model by focusing on the 23 essential competencies of teachers, providing a clearer and more efficient vehicle for evaluating the co-development of student achievement and teacher professionalization. Nevertheless, the Causal Model and the Focused Model have their own strengths and complement each other in the practice of teacher assessment.

3.3 CTEM Model: The Causal Model of Marzano's Teacher Evaluation

. There are four domains in CTEM Model, namely, 'Classroom Strategies and Behaviors', 'Planning and Preparation', 'Reflecting on Teaching', and 'Collegiality and Professionalism', each domain contains 2-3 dimensions, for a total of 11 dimensions covering 60 elements (as shown in Figure 1) (Charlotte Danielson 2013).

Classroom strategies and behaviours feature are prominent in CTEM, a section where students are truly engaged in content learning. It assesses three main components of classroom teaching: procedures, content and practice, and contains 41 elements that focus on teaching strategies directly related to student achievement. The Planning and Preparation section demonstrates the rationality of course planning and teaching design, and is divided into eight elements along the three dimensions. The reflective teaching section illustrates teachers' reflections on their teaching practice and pursuit of personal professional development. Reflective teachers are skilled at gathering professional and informative feedback on their teaching and sharpening their professional skills in order to address the challenges of improving student achievement. The area of collegiality and professionalism is not directly linked to teaching and learning activities, but provides a good environment for effective implementation of each area. Specialist teachers make full use of every platform to improve their professional skills. Therefore, educational authorities, school administrators and relevant staff should be actively involved in the creation of collaborative communities.

Table 2. Marzano Teacher Evaluation Causal Model Evaluation Criteria

Classification	**The usefulness of 60 elements**
Non-effective	Strategies are proposed but not applied.
Beginning	Not fully applying the strategy correctly or somehow lacking strategies are proposed but not applied.
Developmental	Have clear learning objectives that describe the student's performance.
Applied	Have clear learning objectives that describe the student's level of performance and monitor their performance.
Innovative	Develop new strategies that are relevant and meet the needs of students

Similar to the Danielson Framework for Teaching and Learning, the Marzano Teacher Evaluation Model has its own evaluation criteria, which is divided into five levels: non-effective, beginning, developmental, applied and innovative (as shown in Table 2).

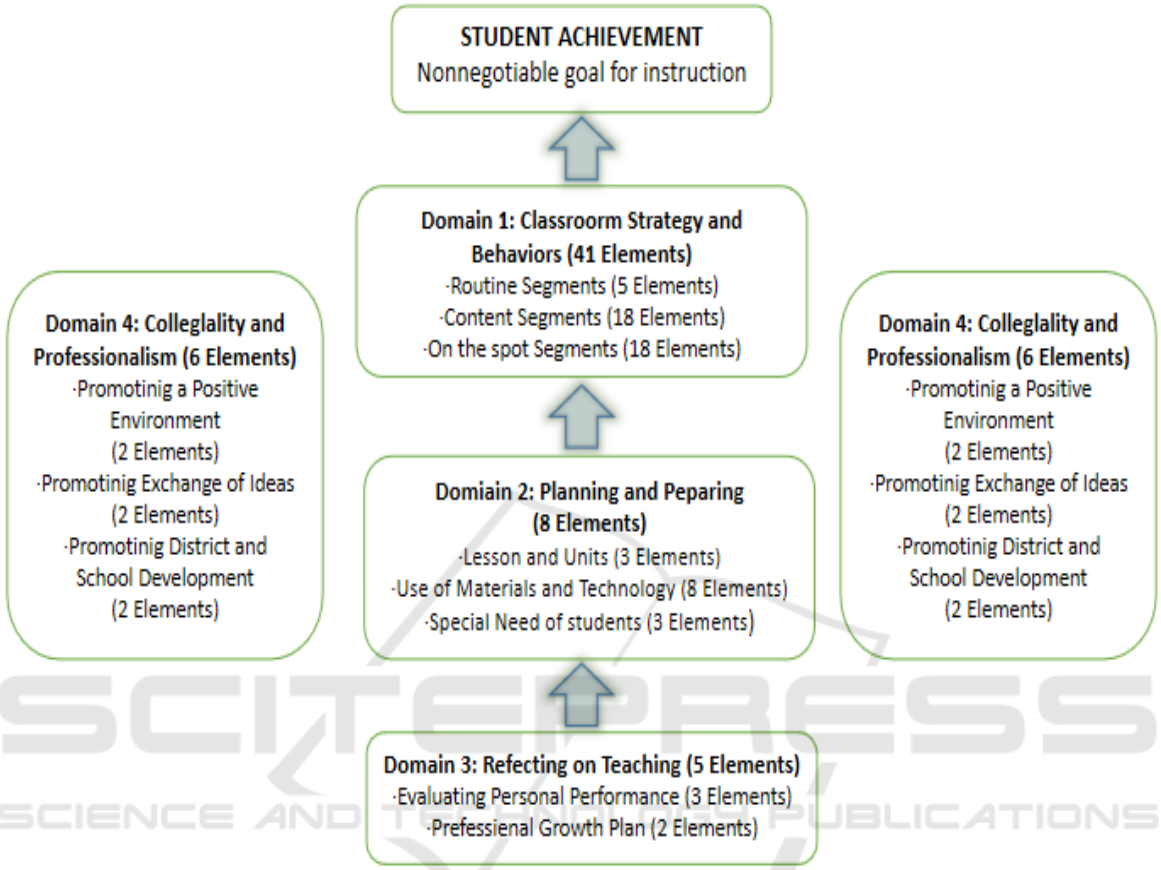

Figure 1. Marzano Causal Teacher Evaluation Model

3.4 FTEM Model: The Focused Model of Marzano's Teacher Evaluation

The Marzano Focused Teacher Evaluation Model differs from the original model in that it reduces the number of teacher skills that affect student achievement from 60 to 23, emphasizing the assessment's efficiency. For any of the 23 basic teacher competencies, FTEM Model carefully designs them in indicator descriptions and desired outcomes. As shown in Figure 2 , the 23 core teacher competencies are concentrated in four areas: 'Standards-based planning', 'Standards-based instruction', 'Conditions for learning' and 'Professional responsibilities' (Lv, 2015). Domain 1 (Standards-based planning) is a consolidation of the eight elements of 'Planning and preparation' in Casual Model. Domain 2 (Standards-Based Instruction) is the core observation of the model, encompassing 10 areas of student learning from basic to cognitive improvement, and significantly reducing the 41 elements of the Causal Model of 'classroom strategies and behaviors'. Domain 3 (Conditions for learning) is a complementary condition to Domain 2 (Standards-based instruction) and provides support for effective teaching and learning. Domain 4 (Professional responsibility) is largely aligned with the Causal Model's 'collegiality and professionalism'. To some extent, the Focus Model is an inheritance and development of the Causal Model.

The Focused Model differs from the Causal Model in seven ways: it is standards-based, emphasizing that the development of standards should follow the dynamics of teaching and learning; it is centralized and simplified, focusing on the assessment of 23 core teacher competencies; it is transparent, providing objective feedback on evaluation results; it is authentic, with tests conducted with

teachers from hundreds of schools and districts to collect data; it is transparent, providing objective feedback on evaluation results; and it is authentic, with tests conducted with teachers from hundreds of schools.

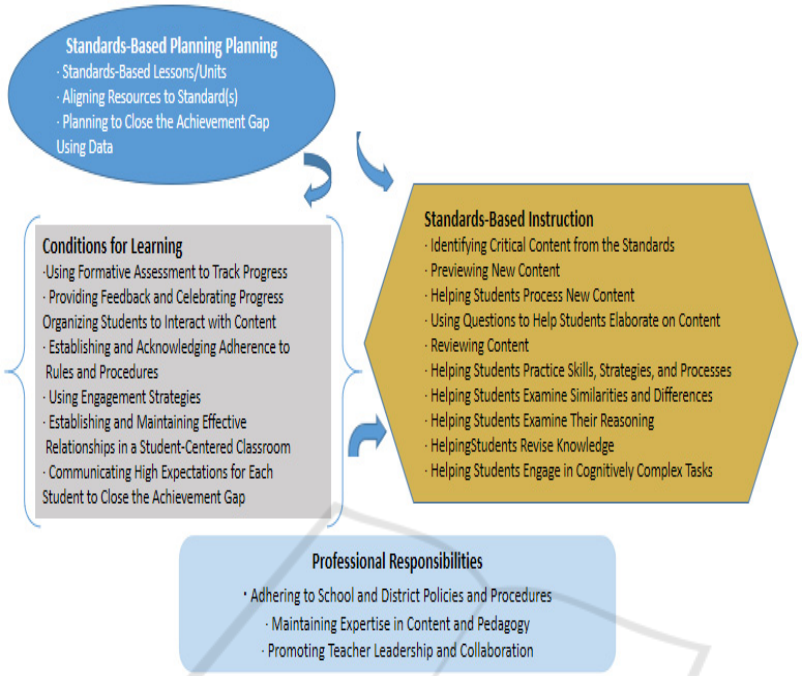

Figure 2. Marzano Focused Teacher Evaluation Model

4 COMPARISON OF THE SIMILARITIES AND DIFFERENCES BETWEEN TWO MULTIPLE TEACHER EVALUATION MODELS

4.1 Similarities Between the Two Teacher Evaluation Models

4.1.1 Evaluation Based on Multiple Evidence

In the process of evaluating teachers' teaching activities, teachers should not only have 'process texts' that demonstrate the unit or lesson plans they have designed, assess students' basic classroom performance, but also have 'explanatory texts' to examine and reflect on their teaching activities, engage in teacher-student interaction and peer communication. The former shows the evidence of teacher's teaching activities and the latter presents the evidence of teacher's professional development. In terms of these two models, the Danielson Teacher Evaluation Model is all about collecting multiple sources of evidence, which come from the school, the teacher, and students. These evidence are then mapped to the elemental descriptions and specific examples of each dimension in the model. The Marzano Focused Model is also evidence-oriented, which collects hundreds of feedback from various schools to form the model. The above shows that the US multiple teacher evaluation is not a generalized description of evaluation, but is based on multiple evidence of the evaluation process.

4.1.2 Focus on Knowledge, Skills and Professional Quality Aspects

. From the evaluation indicators in the two main models, it is clear that the content can be divided into three points: professional knowledge (should know), professional skills or practices (can do) and professional qualities (willing to hold). Practical knowledge is the basis for teachers to become teachers. Practical skills are the ability of teachers to apply their prior practical knowledge in the classroom, with students of different personalities, potentials

and home backgrounds. Professional qualities are concerned with teachers' reflective self-awareness, professional judgement and educational wisdom. In short, these two US multiple teacher evaluation models can generally be considered internally consistent in content.

4.1.3 Development-Oriented Values

In terms of the values advocated by the Danielson Model and the Marzano Model, it is clear that the US multiple teacher evaluation model is no longer seen as a tool for judging teachers, but in a large sense, is focused on the 'development' of the 'person', i.e. the 'professional development' of the teacher, the personal development of students.The Danielson Framework for Teaching has been dedicated to teacher professional development since its establishment by establishing criteria for evaluating instructors' classroom teaching. Meanwhile, the Marzano Teacher Evaluation Model employs innovative educational assessment to investigate the link between teacher effectiveness and student accomplishment, focusing not only on teacher evaluation but also on student academic advancement. Admittedly, two evaluation models both use their own scientific and feasible systems to propose practical paths for teachers' professional growth.

4.2 Differences Between the Two Models of Teacher Evaluation

4.2.1 Opportunity: Contingency and Permanence

The Danielson Framework for Teaching is based on the teacher accreditation system in the United States. Danielson worked on the Praxis III in order to compensate for the deficiencies of the Praxis I and Praxis II assessments. He believed that the Praxis III would be useful not just for diagnosing teachers' classroom teaching standards, but also for encouraging teachers' professional development by assessing their teaching practice. As a result, the Danielson Framework for Teaching's birth was fortunate. On the other hand, the Marzano Teacher Evaluation Model was the outcome of a confluence of three variables. Specifically, it was a combination of the environment of teacher evaluation reform in the United States, the emphasis of authorities on educational accountability and the rise of multiple evaluation systems, and the accumulation of Marzano's years of research. It is clear that the Danielson Framework and the Marzano Model are significantly different.

4.2.2 The Model's Philosophy: Teacher Subject and Teacher-Student Growth

.Although the Danielson Framework and the Marzano Model both emphasize teachers' professional development, however, their philosophy are slightly different. The former focuses more on teacher subject, especially how teachers transfer from novices to experts. While the latter seeks the joint growth of teachers' professional development and students' academic achievement. Specifically, the first three domains of the Marzano Focused Model stress on grading student evidence of learning to reduce pressure on teachers to "perform" , and shift the key point to student attainment of standards; the last domain is about teacher development.

4.2.3 Model Framework: Traditional Criteria and Factor Focus

.Through the comparison of the two models, it is obvious that there are also differences between them in operation. The Danielson Framework provides teacher users with a practical reference--a framework containing board-dimension-point index, giving detailed descriptions on the points corresponding to 22 dimensions of the four boards. The core of the Marzano Teacher Evaluation Model, however, is to explore scientifically the correlation between teacher effectiveness and students achievement, and to distil the key areas and elements that influence them. As the initial Causal Model covered four domains and 60 elements, the subsequent Focused Model corrected the ambiguity and redundancy of the model in practice, simplified the previous elements into 23 essential teacher competencies in an effort to make the assessment process efficient and clear. Thus, it can be concluded that for teacher users, Danielson Framework is fixed and rigorous, while Marzano Model is multi-angle and flexible.

5 CONCLUSIONS

At present, the overall research on teacher evaluation in China is still in its early stages, especially the research on teacher evaluation practice is radically weak. Through a comparative analysis of the background and evaluation framework of two teacher evaluation models, the following insights are pro-

vided for the current teacher evaluation practice process in China.

To begin, teacher evaluation should focus on the evaluation of classroom teaching practice, with the ultimate goal of improving “teaching” and “learning”. Since the 21st century, the value orientation of teacher evaluation practice models has been tightly focused on two real needs: the need for teachers to improve their professional development, and the need for teachers' public accountability. As a result, the field of teacher evaluation has made it a top focus. China's Outline of Basic Education Curriculum Reform also specifies that teacher evaluation in the new curriculum reform should focus on teachers' classroom teaching process and continuously enhance teachers' teaching standards.

Second, selecting and integrating teacher evaluation models is critical, model suitability may be a high concern. Whether it's a Chinese-developed teacher evaluation model or a classical model imported from outside, the applicability of the model must be examined. Three main components are considered when selecting and integrating models. The first section consists of a summary and refinement of the common traits. The second component is to look at the educational policy environment from the perspective of teacher modes that are built using the country's present evaluation standards, as well as the ability to monitor the educational policy environment for mutual reinforcement. The next step is to take a micro viewpoint, in which the model is chosen and developed depending on the specific realities of a certain school district or school.

Finally, the framework of teacher evaluation indicators should be based on teacher professional standards, which include knowledge, abilities, and attributes. Because the Danielson Framework and the Marzano Model are constructed around these three qualities, we should make sure that our teacher evaluation indicators do as well. Furthermore, the teacher evaluation model can highlight the significance of indicator precision in enhancing evaluation efficacy. As a result, teacher evaluation indicators in China should be based on the idea that the three dimensions of teacher professional standards should be continuously enhanced, and that redundancy should be avoided in the pursuit of precision.

REFERENCES

Charlotte, Danielson. Enhance professional practice: A Framework for Teaching [M]. Virginia: Association for Supervision and Curriculum Development Alexandria, 2013.

Charlotte Danielson. The Framework for Teaching Evaluation Instrument (2013 Edition) [EB/OL]. http://www.teachscape.com/products/focus,2013,pp12-113.

Hursh, David. Neo-Liberalism, Markets and Accountability: Transforming Education and Undermining Democracy in the United States and England. Policy Futures in Education, 2005, pp.3-15.

Hunzicker, Jana. Using Danielson's Framework to Develop Teacher Leaders. Kappa Delta Pi Record, 2017, pp.12-17.

Larsen, Marianne A. A Critical Analysis of Teacher Evaluation Policy Trends. Australian Journal of Education, 2005, pp.292-305.

Lu Ruping. Research on the Construction and Application of Danielson Teacher Evaluation System. Shanghai Normal University, 2006.

Lv, Minxia. A Practical Exploration of Improving the Professional Evaluative Power of Instructional Evaluators in the United States. Comparative Education Research, 2015, pp.72-78.

Marzano Center. 2017 Update: The Marzano Focused Teacher Evaluation Model, [EB/OL]. https://www.learningsciences.com/

Marzano Center. Robert J. Marzano Archives, [EB/OL] https://www.learningsciences.com/authors/robert-j-marzano.

Robert J., Marzano. The Marzano Teacher Evaluation Model [M]. Marzaon Research Library, 2013.

Study, [EB/OL]. http://www.metproject.org/downloads/MET_Ensuring_Fair_and_Reliable_Measures_Practioner_Brief.pdf

Wang Jian. A philosophical examination of the history of the development of educational evaluation. Research in Teacher Education, June, 2008, pp.1-6.

Wu Xiaopeng, Wu Xiaoxia, and Zhang Yi. Constructs and mechanisms of teacher teaching assessment systems. Comparative Education Research, September, 2019.

Zhou Jiushi. Introduction and Discussion on A Framework for Teaching. Basic Education, 2017, pp.88-94.