

Review

Sustainability of Evaluation: The Origin and Development of Value-Added Evaluation from the Global Perspective

Xiaopeng Wu ^{1,2} , Tianshu Xu ³ and Jincheng Zhou ^{2,4,5,*} 

¹ Faculty of Education, Northeast Normal University, Changchun 130024, China

² School of Computer and Information, Qiannan Normal University for Nationalities, Duyun 558000, China

³ College of Teacher Education, East China Normal University, Shanghai 200062, China

⁴ Key Laboratory of Complex Systems and Intelligent Optimization of Guizhou Province, Duyun 558000, China

⁵ Key Laboratory of Complex Systems and Intelligent Optimization of Qiannan, Duyun 558000, China

* Correspondence: zjc81@sgmtu.edu.cn

Abstract: Education evaluation plays a key role in promoting education development. The sustainable concept of evaluation provides the basis for the sustainable development of education. Value-added evaluation makes up for the shortcomings of traditional evaluation that only focuses on the results. It takes the development of students and teachers and the improvement of the education system as the main variables of evaluation, providing a basis for the sustainable development of students. This study summarizes the origin and development of value-added evaluation, including its theoretical basis, value orientation, evaluation content and typical cases, and attempts to gain a deeper understanding of it through multiple evaluation methods. The research shows that the value-added evaluation showed a trend of more diversified evaluation indicators, diagnostic evaluation results, and emphasis on longitudinal analysis; value-added evaluation is based on the relative increase in value and emphasizes the “net increment” of students’ learning achievements; the content of value-added evaluation focuses on students’ academic achievements and teacher effect; the evaluation methods mainly include direct evaluation method, indirect investigation method and multivariate and hierarchical statistical method. This research has carried out a comprehensive analysis and interpretation of value-added evaluation to ensure the deep understanding and rational application of it.

Keywords: value-added evaluation; education evaluation; value-added; education quality; education equity



Citation: Wu, X.; Xu, T.; Zhou, J. Sustainability of Evaluation: The Origin and Development of Value-Added Evaluation from the Global Perspective. *Sustainability* **2022**, *14*, 15867. <https://doi.org/10.3390/su142315867>

Academic Editor: Yuto Kitamura

Received: 19 October 2022

Accepted: 27 November 2022

Published: 29 November 2022

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The education evaluation was coeval with the education accountability and the educational accountability also promoted the educational equity. At the beginning of the 20th century, with the expansion of educational resources and investment in Europe and the United States, discussions on promoting educational reform and improving educational effectiveness became a hot spot. In 1966, the United States released the education quality report—Coleman Report. The report pointed out that the key factor to determine the quality of education in a school is whether the school has effectively helped students with the problem of educational equity caused by birth. The efforts made by the school to help students grow should be the main factor in educational evaluation. This view has formed a new system to evaluate education quality based on students’ progress, which has become the basis of value-added evaluation, that is, to investigate the value-added role of schools on students [1]. The term “value-added” originated from economics and refers to the difference between input and output. The economists who were interested in education applied the input and output model to educational evaluation to understand how various factors affect educational outcomes [2]. In 1983, the United States promulgated

The Comprehensive Education Reform Act (CERA), which proposed a teacher effect evaluation system. On this basis, William L. Sanders and McLean A. Robert, two statisticians from the University of Tennessee, and their team applied the method of statistical mixed model to the educational statistical evaluation system [3]. In 1984, William Sanders and others first proposed to use value-added scores to evaluate school and teacher effect, using student achievement data as the basis for teacher evaluation [4]. In 1985, Terry Taylor explicitly put forward the value-added evaluation method for the first time, using the term “value-added” to express students’ learning progress and development embodied in the learning process [5].

In the 1970s, value-added assessment in education gradually came into force in the world. By the end of the 1980s, value-added evaluation had been widely used in educational practice and became an important basis for countries to formulate educational assessment policies and evaluate the quality of school education [6]. In 1987, Ernest L. Boyer applied the concept of value-added evaluation to the evaluation of students’ achievements. Over the same period, the Truman State University used value-added evaluation to evaluate students’ basic knowledge and skills and evaluated major disciplines in the form of national standard examination [7]. Value-added evaluation has not only developed rapidly in the United States, but also been widely used in some European countries. In the UK, value-added evaluation has been piloted since the 1990s. In 1993, Ron Dearing suggested that the results of the school evaluation should use value-added information as an important indicator of school accountability and effect evaluation [8]. It was extended to England and Wales in 2002 and widely used nationwide in 2006. With the continuous development of educational technology and evaluation methods, value-added evaluation has become the main measure of accountability of British schools and the guarantee of education quality [9]. The assessment of school progress in the UK has changed from “value-added” orientation to “contextual value-added” orientation [10].

Since the 21st century, value-added evaluation has received widespread attention in various countries and organizations and diversified attempts of it have been performed. At the beginning of the 21st century, Canada included value-added evaluation as the main basis for school effect evaluation. In the form of “added value”, France has included value-added evaluation in the index system of the prediction and evaluation organization of the Ministry of Education for the evaluation of schools. The Organization for Economic Cooperation and Development (OECD) has also included value-added evaluation into the indicator project of international education assessment, so as to enhance the effectiveness of international education comparison [11]. Since 2000, Hong Kong has also established the School Value-Added Information System (SVAIS), which evaluates schools in terms of the increase in students’ academic performance, social performance, and emotional indicators. The emergence of this system has played an important role in improving the quality of school teaching and promoting students’ academic development in Hong Kong. On June 30 2020, the General Plan for Deepening the Reform of Educational Evaluation in the New Era was deliberated and adopted at the meeting of the State Council of China, which clearly proposed to explore value-added evaluation and establish a scientific educational evaluation system and mechanism that meet the requirements of the times [12]. In the UK, the USA, and elsewhere, school accountability systems increasingly compare schools by using students’ scores in high-stakes standardized tests as value-added measures of school performance [13].

Value-added evaluation has been applied in all aspects of education evaluation, ranging from teachers’ teaching qualities to the effect of the entire education system. In the development process of value-added evaluation, we can draw three trends: ① The evaluation indicators have changed from being single to more diversified. With the improvement of evaluation technology and the complexity of the model, evaluation indicators can be multivariable and multi-level, rather than just based on students’ academic achievements at first. ② The evaluation results have gradually changed from summative to diagnostic. At first, value-added evaluation was mainly used in educational accountability, but with

the continuous development of formative evaluation, value-added evaluation has paid more attention to the diagnostic function of the evaluation process, providing process data for the improvement of education. ③ The perspective of evaluation has changed from horizontal comparison to longitudinal analysis. At first, the main function of the value-added evaluation was to compare and rank the effect of the education system and institutions and then determine the investment of funds and the allocation of educational resources. In recent years, the value-added evaluation tends to focus on the diagnosis of possible problems through the analysis of longitudinal data, so as to support educational decision-making. In short, large-scale assessment has become one of the foundations of accountability-based systems and policies, not only for regions, schools, and students but also for teachers [14].

2. The Theoretical Basis and Value Orientation of Value-Added Evaluation

Any evaluation activity originates from a deep understanding of things. The motivation of evaluation often comes from the comparison and even sequencing of different evaluation subjects. However, the ultimate goal of evaluation should be to promote the continuous development of the subject while deeply understanding the evaluation subject, as is the case with education evaluation. In the view of Ralph Taylor, the founder of modern educational evaluation, educational evaluation is a process of comparison between the established goals and the actual results. Value-added evaluation originated from the “Input Environment Output” (IEO) model. The purpose of this model proposed by Astin W. Alexander is to separate students’ background factors from the environment. He believes that the quality of higher education is the result of the joint action of “input” (education investment, students’ experience, and characteristics) and “environment” (students’ atmosphere and social interaction), and “output” in the way of students’ achievement [15]. The IEO model is the premise of value-added evaluation. Similarly, Pace C. Robert proposed the concept of the “Quality of Student Effort” (QSE) for students’ “value added”. He believed that students’ efforts were determined by their time and effort in class. The more practice and effort they put in, the greater their effort [16]. These theories have become the basis for the development of value-added evaluation.

Among the theories of educational value-added evaluation, Alexander Austin is undoubtedly the most important person, the most famous theory of whom is the theory of interaction between individual and environment, based on which the IEO model was proposed. In addition, in the late 1970s and early 1980s, he put forward the famous Student Involvement Theory by systematically summarizing decades of large-scale longitudinal research data, which is similar to Pace’s QSE theory, believing that the more time and energy students put into meaningful activities, the greater their gains in school experience [17]. The standard to measure the quality of education is the degree to which students effectively participate in school activities. The ideas of “participation”, “process”, and “all-round development” in this theory play a crucial role in the generation and development of value-added evaluation [18]. In 2006, on the basis of student participation theory, George D. Kuh put forward the concept of student engagement, that is, the degree of student participation in activities. The concept of student engagement is composed of two elements: one is the time and energy for students to participate in academic activities of the school and the other is the creation of learning resources, learning opportunities, and environment that the school can provide for students [19].

In the development of value-added evaluation, The Theory of School Effect also plays an important role. Is school education effective? How to evaluate the effect of the school? Around these problems, the theory of school effectiveness came into being. The United Nations Educational, Scientific, and Cultural Organization (UNESCO) once expressed the school effectiveness as a significant role in students’ academic achievements, excluding the school background and community environmental factors, that is, the role in improving students’ reading, writing, academic, and social skills through teaching practice, regular activity organization, and management. In addition to these theories, cognitive structure

theory, social psychology and identity development theory, type theory, integration theory, etc. have played a certain role in the development of value-added evaluation [20].

Value-added evaluation is based on “education value added”, which emphasizes the progress achieved by students in a period of time [21]. The value orientation of value-added evaluation is to pursue the maximization of gains, the premise of which is to recognize that all students have the right to a fair education and progress. Excellent schools should be able to promote students’ development and have a positive impact on their lives [22]. Value-added evaluation emphasizes the “added value” or “increment” in the process of education development, which can be the improvement of the school’s teaching quality, students’ overall quality, or the comprehensive capability of school [23].

Harvey and Green believed that the increment is a measure of quality and it is the promotion degree of knowledge, skills, and abilities that education brings to students [24]. In a word, the value orientation of value-added evaluation lies in the increase in efficiency, whether it is the increase in students’ learning achievements or the increase in school efficiency, emphasizing the relative increment. This concept introduced from economics has enhanced the efficiency of education and brought an impetus to the development of education. However, this value orientation also tends to reveal problems, linking the education of people with the production efficiency of products and causing the way of educating people to be more inflexible. Therefore, the increment of education is not only efficiency but also the cultivation of students’ character and the reinforcement of culture, which is difficult to examine with pure increment. Therefore, value-added evaluation should also adhere to the humanistic value orientation and put human development first.

3. Evaluation Contents and Typical Cases of Value-Added Evaluation

3.1. Academic Achievement Assessment

The assessment of students’ academic achievements is mainly concentrated in the elementary education stage, but the value-added evaluation of students is basically in the service of accountability. Using longitudinal data to evaluate students’ achievements is one of the most important evaluation contents of value-added evaluation. The core content of the No Child Left Behind Act of 2002 is to use student achievement data to evaluate the school effect [25]. Driven by the Act, many states have established student evaluation data systems, which provide students’ academic performance data for teachers’ professional development [26]. In the assessment of students’ academic achievements, the Pupil Achievement Tracker (PAT) in England shows the method and process of academic achievement assessment [27]. Value-added evaluation refers to the progress of students’ academic achievements in the context of England. The process of evaluation is to analyze the improvement of students’ academic achievements [28]. The characteristic of value-added evaluation is to change from static absolute performance measurement to dynamic relative performance improvement. During the evaluation process, the increment is calculated by the difference between the grades of senior and junior grades. In the value-added assessment of elementary education in England, students in primary stage (PS) and secondary stage (SS) are selected for assessment and the value-added PS1-2 in the primary stage and value-added SS2-4 in secondary stage are calculated, respectively. The value added of PS1-2 refers to the growth rate of students in PS1 and PS2. The test subjects are mathematics, writing, and reading. The difference between this assessment and the traditional test is that it emphasizes the increase in students’ scores in a period of time, which belongs to a dynamic assessment and excludes the factors of students’ initial capability. The calculation method of SS2-4 value added is similar to that of PS1-2, but SS2-4 increases the number of subjects to eight. In order to highlight the importance of mathematics and English, the scores of these two subjects are double weighted [9]. The value-added calculation method of SS2-4 is:

$$Y_{SS2-4} = \frac{\sum_{i=1}^2 2 \cdot x_i + \sum_{j=1}^6 x_j}{10} \quad (1)$$

In (1), i represents the primary school and j represents the middle school.

In recent years, the research focus of teacher education has gradually shifted from teacher effects to teachers' contributions to students' academic performances and the evaluation of students' academic performance has become a focal point [29]. In addition, research has investigated the academic achievements of students in 24 primary schools in the UK. The value-added effects of the schools in the sample on students' growth are evaluated using a simple fixed-effect model. The results of this study show that there are significant inconsistencies between the ranking based on the value-added effect of schools on student progress and the ranking based on the average student achievement [30].

With the continuous development of modern statistical models, the evaluation of students' academic achievements are more systematic. Decision trees [31,32], multiple linear models [33,34], etc. can be used to measure students' value added in a multivariate and hierarchical way, which is more operable to promote students' development.

3.2. Evaluation of Teachers' Effects

The value-added evaluation of teachers' effects comes from the consideration of factors other than teachers' teaching. The senior teachers usually have better opportunities to choose teaching resources, teaching environments, high-quality students, and other factors [35]. The traditional way of evaluating teachers only based on students' achievements may exaggerate the teaching effect of these teachers [36]. The value-added evaluation is based on the students' scores in many tests. Modern measurement models such as the multi-layer linear model, multiple linear regression, and decision tree are used to exclude the initial capability of students, school resources, students' family economic status, and other factors that are not under the control of teachers, so that the examination variables can be directly attributed to the professional efforts of educators, thus making the evaluation more objective [37]. In Race to the Top, the US federal government pointed out that highly effective teachers can greatly improve students' academic performances, which is reflected in the change of two standardized tests [38]. In 1992, the state government signed the Education Improvement Act (EIA) and listed the value-added evaluation system as a part of the education accountability system in the Act [39]. The Tennessee Value-Added Assessment System (TVAAS) was established under this background. As the most influential value-added assessment model at present, TVAAS has performed pioneering work in the evaluation of teachers' effects [40]. The TVAAS interprets the teacher effect as "the most important factor affecting the academic growth of student groups is the difference in the effectiveness of individual classroom teachers [39]", that is, "teacher effect is the decisive factor for students' learning differences [41]". The calculation of the teacher effect in the TVAAS is a complex process, which combines the estimation of the average effect of each school system and the average effect of each teacher's students relative to the school system. The evaluation system consists of three models [42]: ① the model used to estimate the average performance of different grades and disciplines in a particular school system each year; ② the model for evaluating the average performance of a particular school; and ③ the teacher model used to estimate the average student performance associated with a particular teacher in the system. The TVAAS is regarded as an accountability revolution, which can provide an objective answer to the problem of teacher effect [43,44]. The TVAAS is not simply an evaluation scheme for teacher effect. It involves a comprehensive evaluation system of students' performances, students' background factors, teacher effects, school effectiveness, and other factors. Under the influence of this system, North Carolina [45,46], Texas, Florida, Louisiana, Ohio, and other states of the United States have incorporated value-added evaluation into the teacher effect accountability system [47,48]. Houston even uses value-added evaluation as the basis for determining teachers' performance salary and dismissing teachers [49]. It is thus clear that at the level of institutions of higher learning and the education system, more attention is paid to the performance evaluation and accountability of teachers and less attention is paid to the value-added evaluation of

students' academic achievements, which is closely related to the functions of institutions to a large extent.

4. Method of Value-Added Evaluation

Value-added evaluation is a kind of educational production function model. Different from the ordinary educational input–output model, value-added evaluation separates students' academic achievements from other interfering variables (such as the quality of students, gender, nationality, etc.) through relevant statistical analysis techniques then examines the degree of students' academic progress in a certain period and uses this “increment” to measure whether students, teachers, and schools have achieved relatively greater development [50].

In the analysis process, different statistical models and psychometric models are often used in combination with specific tests to support the educational accountability system through the collection, analysis, management, and maintenance of longitudinal data and provides information for educational decision-making and accountability [40]. Although there are differences in the field or object of value-added evaluation, some focus on the academic performance of primary and secondary students, some focus on the resource allocation of the education system, and some focus on the accountability of the higher education system. From the perspective of evaluation methods, value-added evaluation can be roughly divided into three types, namely, the direct evaluation method, indirect survey method, and multivariate and hierarchical statistical method.

4.1. Direct Evaluation Method

The direct evaluation method generally uses the average of students' original scores as a value-added indicator and analyzes the value added of data through descriptive statistics such as average, median, mode, variance, etc. It can be used to directly measure students' learning effectiveness and the measurement results can also be used for many purposes, such as solving educational problems, policy improvement, and teacher evaluation [5]. The “Collegiate Learning Assessment (CLA)” jointly developed by RAND's Value-added Assessment Initiative (VAAI) in the United States is a typical example of direct assessment. The assessment samples freshmen for testing and retests the same group of students in their senior year to compare the growth of students' ability test scores in the past four years. The assessment covers four aspects: critical thinking, collaborative communication, demonstration analysis, and problem solving. The test lasts for 90 min. So far, more than 500 universities and 250,000 students have participated in the test [51]. In addition, there are also examples such as the Collegiate Assessment of Academic Proficiency (CAAP), Measure of Academic Proficiency and Progress (MAPP), Voluntary System of Accountability (VSA) [52], and the Graduate Skills Assessment (GSA) in Australia [53]. Although the direct measurement method has been widely used in higher education, some studies believe that the direct measurement method is more suitable for elementary and secondary education. Because the higher education system is relatively complex—the increment that affects students' achievements may be composed of a variety of factors, such as academic ability, personality traits, social influence, etc.—the changes in students' achievements cannot be directly attributed to schools or teachers. This kind of standardized direct evaluation method has also developed from simply testing students' mastery of learning content to the comprehensive evaluation of students' ability levels [37].

4.2. Indirect Survey Method

The indirect survey method does not directly obtain the added value of students' academic achievements through pre-post-tests but infers the progress of students by investigating students' learning behaviors, learning experiences, and teachers' teaching behaviors. The indirect survey method is usually carried out in the form of self-reporting. The student academic achievement questionnaire also belongs to this form. The survey content includes curriculum learning, community activities, teacher–student interaction,

classmate communication, condition utilization, resource acquisition, etc. [30]. The National Survey of Student Engagement (NSSE) organized by the Pew Charitable Trusts (PCT) is a typical indirect survey, which aims to improve the learning quality of university students. The respondents of the survey are senior students, the survey content includes five indicators: students' challenges, the teacher–student interaction level, active cooperative learning level, support of campus environment, and richness of teaching experience. A total of 42 aspects of undergraduate experience were collected; 1451 colleges and universities participated in the survey, with more than 2 million people [54]. In addition, there are also the National Survey for Student Engagement (NSSE), College Student Experience Questionnaire (CSEQ), Course Experience Questionnaire (CEQ), the monitoring research on teaching quality and student development of capital universities developed by Peking University and the investigation on learning experience of Chinese university courses organized by Beijing Normal University, Australia's National Assessment of Primary, and Secondary School Effectiveness in Five Areas [55]. This method fully considers the covariates that may affect the value added of education, such as students' backgrounds, students' experiences, and school characteristics. It provides rich information in understanding students' own advantages and disadvantages, which is conducive to students' self-improvement, supports decision-making, and effectively responds to social accountability [5]. However, it is believed that the indirect survey method has such problems as information accuracy [56], effective response rates [57], halo effect [58,59], and cannot effectively replace the objective test to measure students' ability and growth increment. Therefore, in the indirect survey method, more types of data can be considered to enhance the accuracy of the survey.

4.3. Multivariate and Hierarchical Statistical Method

Educational statisticians use complex statistical models to nest multi-level data such as students, teachers, and schools in the educational structure environment to obtain the educational increment. The research suggests that the traditional VAM estimation is biased, which has promoted the development of a hierarchical model describing the joint distribution of school value-added, bias, and lottery compliance. This has also attracted attention to the multi-layer model [60]. Generally, there are two models, namely, the Multivariate Model and Hierarchical Model. The multiple linear regression model is the most typical example of Multivariate Models, which are used to analyze the relationship between a dependent variable and multiple independent variables. The actual growth of a students' performance can be obtained according to the difference between the observed value and the expected value. The residual value obtained through multiple linear regression is the increase in students' learning progress over a period. This method can incorporate various factors affecting students into the statistical model to comprehensively analyze the relationship between internal variables and cross layer variables in the aspect of students and schools, so as to objectively evaluate the importance and interaction of various factors affecting students' achievements [18]. In addition, the cross-classification model is also commonly used in research [61]. The model divides the variation within the class into two parts: systematic variation and random variation. It can well evaluate the cross-grade correlation of students' scores and the cumulative effect of teachers on students [62]. The Hierarchical Linear Model (HLM) [63], as a representative of Hierarchical Models, has been widely used in value-added evaluation. The Hierarchical Linear Model complements the methodology and theory of longitudinal research on student change or research on the structure of students' "nested" school education environment, which are common in education research. This model promotes the research on the effect of each level of covariates in nesting on dependent variables, effectively reduces research errors, and enhances the reliability and validity of the research [64]. The Hierarchical Linear Model separates the external variables that affect students' academic performance, such as learning foundation and family background, from the effects of teachers and schools and obtains the "net effect" of teachers and schools, thus distinguishing the responsibilities of all parties [65]. This method has been widely used in large-scale tests such as the Program for International

Student Assessment, Trends in International Mathematics Science Study, and Progress in International Reading Literacy Study. These research plans not only consider the hierarchical nature of the data, but also show the longitudinal development of the data, fully reflecting the role of value-added in promoting school effect evaluation and education quality, and the guidance of education policy will be more general.

In addition, with the development of evaluation technology, many new technologies and methods have been introduced into value-added evaluation methods, such as random forest [66], hierarchical models based on big data [67], Student Growth Percentile (SGP) models [68], and multiple models (empirical Bayes predictions) based on empirical Bayesian prediction [69]. Of course, there are shortcomings in any model. Some studies have conducted special analysis on the errors and deviations of value-added evaluation models [70–72] and the effectiveness of different value-added models is also questioned [73,74].

5. Conclusions

Value-added evaluation, as an important evaluation method to reasonably evaluate the educational effect, promote educational equity, and maintain the sustainable development of education, plays an important role in stimulating the endogenous force of education. This study summarizes the origin and development of value-added evaluation, analyzes the theoretical basis of value-added evaluation, and further discusses the basic value orientation of value-added evaluation. Based on the case of value-added evaluation, this paper interprets the basic content of value-added evaluation from two aspects: student achievement evaluation and teacher effect evaluation. The evaluation methods of value-added evaluation are analyzed from three aspects: direct evaluation method, indirect investigation method, and multivariate and hierarchical statistical method. It was found that the value-added evaluation showed a trend of more diversified evaluation indicators, diagnostic evaluation results, and emphasis on longitudinal analysis; value-added evaluation is based on the relative increase in value and emphasizes the “net increment” of students’ learning achievements; the content of value-added evaluation focuses on the evaluation of students’ academic achievements and teacher effect; the evaluation methods mainly include the direct evaluation method, indirect investigation method, and multivariate and hierarchical statistical method. Through the above analysis, the following characteristics of value-added evaluation can be further summarized:

5.1. Value-Added Evaluation Focuses on the Starting Point, Distinguishes Responsibilities, and Contributes to Education Equity

The theoretical assumption of value-added evaluation is that school education can add “value” to students’ academic achievements, which is the increase in educational value [11]. Value added refers to the increase in students’ knowledge ability and learning experience after a period of learning, which is the difference of students’ academic gains between the two assessments [5]. This evaluation method can effectively separate the functions of school and teacher effect through modern statistics and measurement technology and examine the teacher effect when excluding background factors such as students’ learning starting point, educational resources, and parents’ background. Therefore, the evaluation results are more objective, which can effectively distinguish the responsibilities of schools, teachers, students and other aspects, and greatly optimize the educational accountability system. Some studies believe that value-added measures are unbiased predictors of teacher performance [75]. This method has also been applied to the formulation of education policies in many countries. The United States reauthorized the Basic and Secondary Education Act in 2002. One of the purposes is to promote the application of value-added evaluation in elementary and secondary education, so as to clarify responsibilities and promote the reform of educational accountability [76]. Many European and American countries have used value-added evaluation as a common strategy to carry out basic education evaluation. For example, the United Kingdom is the main strategy for school accountability, while the

United States focuses on the evaluation of the teacher effect [9], which shows that value-added evaluation plays a huge role in promoting education accountability and education equity. For example, public schools in North Carolina use teacher value-added and evaluation rating data to estimate the signaling and human capital effects of graduate degrees [77]. In addition, research has analyzed the relationship between family income and student value-added [78,79]. Value-added evaluation has accumulated experience of continuous change in the long-term development process, showing good institutional stability.

5.2. Value-Added Evaluation Focuses on Process, Emphasizes Progress, and Provides Full Play to the Diagnostic Function of Evaluation

The process of value-added evaluation is to use specific value added to evaluate the contribution of teachers or schools to students' academic achievements by tracking students' academic changes over a period of time. This evaluation system emphasizes the process of student evaluation and tries to figure out the causes of changes through multiple statistics, decision trees, and other modern measurement methods. The most prominent feature of evaluation is the "value added" of students, so it is an evaluation form with the development of students as the main purpose. In many cases, the school with the best performance measured by the original value is not the school with the best performance measured by the value-added value [80]. Through long-term monitoring data of students, we can further analyze students' learning characteristics, teachers' teaching style, school management mode, and so on, provide full play to the diagnostic function in the evaluation process, and highlight the value-added effect of school education or learning activities on students' learning. With the development of communication technology, it is more practical to use a large amount of data to make decisions on educational data. The diversified analysis of these data can meet various diagnostic needs of education. Value-added evaluation can become mainstream in various countries [81], not only because it has more advanced education concepts but also because it conforms to the characteristics of the times. It combines the needs of summative evaluation for results in education evaluation and the concerns of formative evaluation for processes. The evaluation results are rich in diagnostic information and also have the identification function.

5.3. Value-Added Evaluation Embraces Difference, Evaluates Comprehensively, and Promotes Diversified Development of Various Fields

Value-added models are increasingly used to assess the performance of staff and institutions, from teachers and schools to doctors and CEOs. The effectiveness of the value-added model in performance evaluation mainly depends on the extent to which value-added estimates are biased by selection, such as the differences in the latent abilities of students assigned to teachers [82]. Value-added evaluation is a kind of student-centered developmental evaluation that focuses on the relative development of students. The theoretical basis of value-added evaluation is based on differences. It is believed that different students, teachers, and schools should have different starting points. That is, for the starting point of evaluation, the starting value of these variables is obviously not zero, and their development is affected by the existing basis. Hence, evaluation should be carried out on the premise of considering the difference of actual conditions. Due to the continuous development of evaluation models, the variables that can be considered in value-added evaluation are more complex. Therefore, value-added evaluation is a comprehensive development model with multiple variables. The theory of multiple intelligences believes that human development is affected by multiple intelligence factors and people have their own development in multiple intelligences [83]. The role of education is to discover students' superior intelligence and promote students' diversified development. Value-added evaluation focuses on "increment", which can be the improvement of students' knowledge, skills, emotions, attitudes, and other aspects, as well as the improvement of the teacher and school effect. This greatly promotes students' interest in learning, activates the whole educational accountability system, and pays attention to the progress of the evaluation subject, not just the results. This kind of evaluation considers the interests of all

evaluation subjects, not just the excellent performance of the individual. Therefore, this is an evaluation form that returns to the student standard and emphasizes the all-round development of students [84].

Author Contributions: X.W. constructed the model and completed the manuscript writing; T.X. collected and analyzed the data and corrected the language of the article; J.Z. put forward valuable opinions and provided financial support for the article. All authors have read and agreed to the published version of the manuscript.

Funding: This work was support by 2021 Humanities and Social Science Research project of Ministry of Education of China under Grant No. 21YJC880102; The Top-notch Talent Program of Guizhou province under Grant No. KY [2018]080; the Guizhou Educational Science Planning Project under Grant No. 2021B201; the Qianan Educational Science Planning Project under Grant No. 2021B001; the Guizhou Educational Science Planning Project under Grant No. 2021C046; the Program of Qiannan Normal University for Nationalities under Grant Nos. (qnsy2019005, qnsy2019006, qnsyxk201807, 2021gh19); supported by the Educational Department of Guizhou under Grant KY[2019]067.

Ethics Statements: This study does not contain identifying information, so does not require ethical approval (Committee on Human Research Protection of ECNU. http://life.ecnu.edu.cn/_upload/article/files/de/d9/250ac9ff4d2eb27003ff30bb0269/c4ab83b2-ac10-454d-a352-715b70bb92e4.pdf, accessed on 1 October 2022). No animal studies are presented in this manuscript. No potentially identifiable human images or data are presented in this study.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest: The authors declare that they have no competing interest.

References

1. Zhou, W. Value-added Evaluation: A new choice for evaluating the quality of practical teaching in Higher vocational colleges. *J. Educ. Sci. Hunan Norm. Univ.* **2016**, *15*, 126–128. (In Chinese)
2. Todd, P.E.; Wolpin, K.I. On the specification and estimation of the production function for cognitive achievement. *Econ. J.* **2003**, *113*, 3–33. Available online: <https://www.jstor.org/stable/3590137> (accessed on 1 October 2022). [CrossRef]
3. McLean, R.A.; Sanders, W.L. *Objective Component of Teacher Evaluation: A Feasibility Study (Working Paper No.199)*; University of Tennessee, College of Business Administration: Knoxville, TN, USA, 1984.
4. Sanders, W.L.; Horn, S.P. The Tennessee Value-Added Assessment System (TVAAS): Mixed- Model Methodology in Educational Assessment. *J. Pers. Eval. Educ.* **1994**, *8*, 299–311. [CrossRef]
5. Liu, H. The rise and application of value-added evaluation model in American higher education. *High. Educ. Res.* **2012**, *33*, 96–101. (In Chinese)
6. Bian, Y.; Wang, Y. Value-added evaluation: An effective way to evaluate school running quality. *J. Educ.* **2013**, *9*, 43–48. (In Chinese)
7. Boyer, E.L. *The Undergraduate Experience in America*; Harper & Row: New York, NY, USA, 1987.
8. Dearing, R. *The National Curriculum and Its Assessment: An Interim Report*; National Curriculum Council: New York, NY, USA, 1993.
9. You, J.; Wang, L. The implementation and utility of Value-added evaluation of basic education in England. *Foreign Educ. Res.* **2019**, *46*, 42–53. (In Chinese)
10. Leckie, G.; Goldstein, H. The evolution of school league tables in England 1992–2016: ‘Contextual value-added’, ‘expected progress’ and ‘progress 8’. *Br. Educ. Res. J.* **2017**, *43*, 193–212. [CrossRef]
11. Thomas, S.; Peng, W. Use of “value-added” indicators to assess school performance. *Educ. Res.* **2005**, *9*, 31–35. (In Chinese)
12. China Education Online. The Central Commission for Deep Urban Reform Reviewed and Adopted the Overall Plan for Deepening Education Evaluation Reform in the New Era. 2020. Available online: https://www.eol.cn/news/yaowen/202007/t20200701_1736125.shtml (accessed on 1 October 2022). (In Chinese)
13. Leckie, G.; Goldstein, H. The importance of adjusting for pupil background in school value-added models: A study of Progress 8 and school accountability in England. *Br. Educ. Res. J.* **2019**, *45*, 518–537. [CrossRef]
14. Close, K.; Amrein-Beardsley, A.; Collins, C. *State-Level Assessments and Teacher Evaluation Systems after the Passage of the Every Student Succeeds Act: Some Steps in the Right Direction*; National Education Policy Center: Boulder, CO, USA, 2018.
15. Astin, A.W. *Four Critical Years. Effects of College on Beliefs, Attitudes, and Knowledge*; Josey-Bas: San Francisco, CA, USA, 1977.
16. Pace, C.R. *Achievement and the Quality of Student Effort*; National Commission on Excellence in Education: Washington, DC, USA, 1982. Available online: <https://files.eric.ed.gov/fulltext/ED227101.pdf> (accessed on 1 October 2022).

17. Astin, A.W. *Achieving Educational Excellence: A Critical Assessment of Priorities and Practices in Higher Education*; Josey Bass: San Francisco, CA, USA, 1985.
18. Li, X.; Zhou, Z.; Liang, X. Value-added evaluation and quality assurance in higher education: A review of theories and methods. *Tsinghua J. Educ.* **2013**, *34*, 40–45. (In Chinese) [[CrossRef](#)]
19. Kuh, G.D.; Kinzie, J.L.; Buckley, J.A.; Bridges, B.K.; Hayek, J.C. *What Matters to Student Success: A Review of the Literature*; National Postsecondary Education Cooperative: Washington, DC, USA, 2006; Volume 8.
20. Evans, N.J.; Forney, D.S.; Guido, F.M.; Patton, L.D.; Renn, K.A. *Student Development in College: Theory, Research, and Practice*; John Wiley & Sons: Hoboken, NJ, USA, 2009.
21. Wang, X. Comparison and application of value-added evaluation methods in higher education. *Dev. Eval. High. Educ.* **2018**, *34*, 60–71. (In Chinese)
22. Astin, A.W. *Excellence and Equity in American Education*; ERIC Document Reproduction Service No. ED227098; National Commission on Excellence in Education: Washington, DC, USA, 1982. Available online: <https://files.eric.ed.gov/fulltext/ED227098.pdf> (accessed on 1 October 2022).
23. Cowan, J.; Goldhaber, D.; Theobald, R. Performance Evaluations as a Measure of Teacher Effectiveness When Implementation Differs: Accounting for Variation across Classrooms, Schools, and Districts. *J. Res. Educ. Eff.* **2022**, *15*, 501–531. [[CrossRef](#)]
24. Harvey, L.; Green, D. Defining Quality. *Assess. Eval. High. Educ.* **1993**, *18*, 9–34. [[CrossRef](#)]
25. Crowe, E. *Measuring What Matters: A Stronger Accountability Model for Teacher Education*; Center for American Progress: Washington, DC, USA, 2010. Available online: <https://files.eric.ed.gov/fulltext/ED518518.pdf> (accessed on 1 October 2022).
26. National Research Council. *Preparing Teachers: Building Evidence for Sound policy*; National Academies Press: Washington, DC, USA, 2010.
27. Department for Education, United Kingdom. National Pupil Database: Apply for a Data Extract. 2014. Available online: <https://www.gov.uk/national-pupil-database-apply-for-a-da-ta-extract> (accessed on 1 October 2022).
28. Thomas, S. Value-added measures of school effectiveness in the United Kingdom. *Prospects* **1998**, *28*, 91–108. [[CrossRef](#)]
29. Nye, B.; Konstantopoulos, S.; Hedges, L. How large are teacher effects? *Educ. Eval. Policy Anal.* **2004**, *26*, 237–257. [[CrossRef](#)]
30. Koc, F. Investigation of School Effects on Student Achievement in Primary Education Using Value-Added Assessment. *Eurasian J. Educ. Res.* **2018**, *76*, 41–72. [[CrossRef](#)]
31. Ma, X. Growth in Mathematics Achievement: Analysis with Classification and Regression Trees. *J. Educ. Res.* **2005**, *99*, 78–86. [[CrossRef](#)]
32. Ma, X. Stability of school academic performance across subject areas. *J. Educ. Meas.* **2001**, *38*, 1–18. [[CrossRef](#)]
33. McNeish, D.; Stapleton, L.M.; Silverman, R.D. On the unnecessary ubiquity of hierarchical linear modeling. *Psychol. Methods* **2017**, *22*, 114–140. [[CrossRef](#)]
34. Pollack, B.N. Hierarchical linear modeling and the “unit of analysis” problem: A solution for analyzing responses of intact group members. *Group Dyn. Theory Res. Pract.* **1998**, *2*, 299–312. [[CrossRef](#)]
35. Paige, M.; Amrein-Beardsley, A.; Close, K. Tennessee’s National Impact on teacher Evaluation law & Policy: An assessment of value-added model litigation. *Tenn. J. Law Policy* **2019**, *13*, 523.
36. Braun, H.I. *Using Student Progress to Evaluate Teachers: A Primer on Value-Added Models*; Policy Information Perspective; Educational Testing Service: Princeton, NJ, USA, 2005. Available online: <https://files.eric.ed.gov/fulltext/ED529977.pdf> (accessed on 1 October 2022).
37. Sanders, W.L. Value-added assessment from student achievement data: Opportunities and hurdles. *J. Pers. Eval. Educ.* **2000**, *14*, 329–339. [[CrossRef](#)]
38. McGuinn, P. Stimulating reform: Race to the Top, competitive grants and the Obama education agenda. *Educ. Policy* **2012**, *26*, 136–159. [[CrossRef](#)]
39. Sanders, W.L.; Horn, S.P. Research Findings from the Tennessee Value-Added Assessment System (TVAAS) Database: Implication for Educational Evaluation and Research. *J. Pers. Eval. Educ.* **1998**, *12*, 247–256. [[CrossRef](#)]
40. Sanders, W.L. Value-added assessment. *Sch. Adm.* **1998**, *55*, 24–27.
41. Darling-Hammond, L. Teacher quality and student achievement: A review of state policy evidence. *Educ. Policy Anal. Arch.* **2000**, *8*, 1–44. [[CrossRef](#)]
42. Kupermintz, H. Teacher effects and teacher effectiveness: A validity investigation of the Tennessee Value Added Assessment System. *Educ. Eval. Policy Anal.* **2003**, *25*, 287–298. [[CrossRef](#)]
43. Stone, J.E. Value added assessment: An accountability revolution. In *Better Teachers, Better Schools*; Kanstoroom, M., Finn, C.E., Jr., Eds.; Thomas B. Fordham Foundation: Washington, DC, USA, 1999; pp. 239–249.
44. Grissom, J.A.; Bartanen, B. Potential Race and Gender Biases in High-Stakes Teacher Observations. *J. Policy Anal. Manag.* **2022**, *41*, 131–161. [[CrossRef](#)]
45. Hanselman, P. Do school learning opportunities compound or compensate for background inequalities? Evidence from the case of assignment to effective teachers. *Sociol. Educ.* **2018**, *91*, 132–158. [[CrossRef](#)]
46. Burkhauser, S. How much do school principals matter when it comes to teacher working conditions? *Educ. Eval. Policy Anal.* **2017**, *39*, 126–145. [[CrossRef](#)]
47. Hock, H.; Isenberg, E. Methods for accounting for co-teaching in value-added models. *Stat. Public Policy* **2017**, *4*, 1–11. [[CrossRef](#)]

48. Feng, H.; Rao, C. An Analysis of the Value-added Evaluation of teacher Training Programs in the United States—A Case study of Louisiana. *Comp. Educ. Res.* **2018**, *40*, 59–68. (In Chinese)
49. Corcoran, S.P. *Can Teachers Be Evaluated by Their Students' Test Scores? Should They Be? The Use of Value—Added Measures of Teacher Effectiveness in Policy and Practice*; Annenberg Institute for School Reform at Brown University: Providence, RI, USA, 2010. Available online: <https://files.eric.ed.gov/fulltext/ED522163.pdf> (accessed on 1 October 2022).
50. Scheerens, J.; Glas, C.A.; Thomas, S.M.; Thomas, S. *Educational Evaluation, Assessment, and Monitoring: A Systemic Approach*; Taylor & Francis: Milton Park, UK, 2003; Volume 13.
51. Klein, S.; Benjamin, R.; Shavelson, R.; Bolus, R. The Collegiate Learning Assessment: Facts and Fantasies. *Eval. Rev.* **2007**, *31*, 415–439. [[CrossRef](#)] [[PubMed](#)]
52. Liu, O.L. Measuring Value-added in Higher Education: Conditions and Caveats-Results from Using the Measure of Academic Proficiency and Progress (MAPPTM). *Assess. Eval. High. Educ.* **2011**, *36*, 84–91. [[CrossRef](#)]
53. Rodgers, T. Measuring value added in higher education: Do any of the recent experiences in secondary education in the United Kingdom suggest a way forward? *Qual. Assur. Educ.* **2005**, *13*, 95–106. [[CrossRef](#)]
54. Kuh, G.D. Assessing what really matters to student learning inside the national survey of student engagement. *Chang. Mag. High. Learn.* **2001**, *33*, 10–17. [[CrossRef](#)]
55. Marks, G.N. Should value-added school effects models include student-and school-level covariates? Evidence from Australian population assessment data. *Br. Educ. Res. J.* **2021**, *47*, 181–204. [[CrossRef](#)]
56. Wentland, E.J.; Smith, K.W. *Survey Responses: An Evaluation of Their Validity*; Academic Press: San Diego, CA, USA, 1993.
57. Aaker, J.L.; Williams, P. Empathy versus pride: The influence of emotional appeals across cultures. *J. Consum. Res.* **1998**, *25*, 241–261. [[CrossRef](#)]
58. Bradburn, N.M.; Sudman, S.; Wansink, B. *Asking Questions: The Definitive Guide to Questionnaire Design—For Market Research, Political Polls, and Social and Health Questionnaires*; John Wiley & Sons: Hoboken, NJ, USA, 2004.
59. Kuh, G.D.; Pace, C.R.; Vesper, N. The Development of Process Indicators to Estimate Student Gains Associated with Good Practices in Undergraduate Education. *Res. High. Educ.* **1997**, *38*, 435–454. [[CrossRef](#)]
60. Angrist, J.D.; Hull, P.; Pathak, P.A.; Walters, C. Leveraging lotteries for school value-added: Testing and estimation. *Social Science Electronic Publishing. Q. J. Econ.* **2017**, *2*, 871–919. [[CrossRef](#)]
61. McCaffrey, D.F.; Lockwood, J.R.; Koretz, D.; Louis, T.A.; Hamilton, L. Let's see more empirical studies on value-added modeling of teacher effects: A reply to Raudenbush, Rubin, Stuart and Zanutto, and Reckase. *J. Educ. Behav. Stat.* **2004**, *29*, 139–143. [[CrossRef](#)]
62. Rowan, B.; Correnti, R.; Miller, R. What Large—Scale Survey Research Tells Us about Teacher Effects on Student Achievement: Insights from the Prospects Study of Elementary Schools. *Teach. Coll. Rec.* **2002**, *104*, 1525–1567. [[CrossRef](#)]
63. Bryk, A.S.; Raudenbush, S.W. *Hierarchical Linear Models: Applications and Data Analysis Methods*; Sage Publications: Newbury Park, CA, USA, 1992.
64. Laird, N.M.; Ware, J.H. Random-effects models for longitudinal data. *Biometrics* **1982**, *38*, 963–974. [[CrossRef](#)] [[PubMed](#)]
65. Goldstein, H. Hierarchical data modeling in the social sciences. *J. Educ. Behav. Stat.* **1995**, *20*, 201–204. [[CrossRef](#)]
66. Schiltz, F.; Sestito, P.; Agasisti, T.; De Witte, K. The added value of more accurate predictions for school rankings. *Econ. Educ. Rev.* **2018**, *67*, 207–215. [[CrossRef](#)]
67. Zhang, R. The Application of Big Data in Education Effectiveness Evaluation Based on the Value-Added Approach. *J. Simul.* **2019**, *7*, 57–60.
68. Kurtz, M.D. Value-added and student growth percentile models: What drives differences in estimated classroom effects? *Stat. Public Policy* **2018**, *5*, 1–8. [[CrossRef](#)]
69. Leckie, G. Avoiding bias when estimating the consistency and stability of value-added school effects. *J. Educ. Behav. Stat.* **2018**, *43*, 440–468. [[CrossRef](#)]
70. Kane, M.T. Measurement error and bias in value-added models. *ETS Res. Rep. Ser.* **2017**, *2017*, 1–12. [[CrossRef](#)]
71. Guarino, C.M.; Reckase, M.D.; Wooldridge, J.M. Can value-added measures of teacher performance be trusted? *Educ. Financ. Policy* **2015**, *10*, 117–156. [[CrossRef](#)]
72. Everson, K.C. Value-added modeling and educational accountability: Are we answering the real questions? *Rev. Educ. Res.* **2017**, *87*, 35–70. [[CrossRef](#)]
73. Morris, T.T.; Davies, N.M.; Dorling, D.; Richmond, R.C.; Smith, G.D. Testing the validity of value-added measures of educational progress with genetic data. *Br. Educ. Res. J.* **2018**, *44*, 725–747. [[CrossRef](#)] [[PubMed](#)]
74. Amrein-Beardsley, A.; Close, K. Teacher-level value-added models on trial: Empirical and pragmatic issues of concern across five court cases. *Educ. Policy* **2021**, *35*, 866–907. [[CrossRef](#)]
75. Bacher-Hicks, A.; Chin, M.J.; Kane, T.J.; Staiger, D.O. An experimental evaluation of three teacher quality measures: Value-added, classroom observations, and student surveys. *Econ. Educ. Rev.* **2019**, *73*, 101919. [[CrossRef](#)]
76. Olson, L.; Hoff, D.J. US to Pilot New Gauge of "Growth": Education Department to Permit Shifts in How States Track Gains. *Educ. Week* **2005**, *25*, 1–16.
77. Bastian, K.C. A degree above? The value-added estimates and evaluation ratings of teachers with a graduate degree. *Educ. Financ. Policy* **2019**, *14*, 652–678. [[CrossRef](#)]

-
78. Isenberg, E.; Max, J.; Gleason, P.; Deutsch, J. Do Low-Income Students Have Equal Access to Effective Teachers? *Educ. Eval. Policy Anal.* **2022**, *44*, 234–256. [[CrossRef](#)]
 79. Bau, N.; Das, J. Teacher value added in a low-income country. *Am. Econ. J. Econ. Policy* **2020**, *12*, 62–96. [[CrossRef](#)]
 80. Doris, A.; O'Neill, D.; Sweetman, O. Good schools or good students? The importance of selectivity for school rankings. *Oxf. Rev. Educ.* **2022**, *48*, 806–824. [[CrossRef](#)]
 81. Pascarella, E.T.; Terenzini, P.T. *How College Affects Students: A Third Decade of Research*; Josey-Bas Publisher: San Francisco, CA, USA, 2005.
 82. Chetty, R.; Friedman, J.N.; Rockoff, J. Using lagged outcomes to evaluate bias in value-added models. *Am. Econ. Rev.* **2016**, *106*, 393–399. [[CrossRef](#)]
 83. Gardner, H.E. *Multiple Intelligences: New Horizons in Theory and Practice*; Basic Books: New York, NY, USA, 2008.
 84. Clotfelter, C.T.; Ladd, H.F.; Vigdor, J.L. Teacher Credentials and Student Achievement: Longitudinal Analysis with Student Fixed Effects. *Econ. Educ. Rev.* **2007**, *26*, 673–682. [[CrossRef](#)]