The Double Constraints of Educational Evaluation—A Case Analysis of the College Entrance Examination Reform of China

Weihe Xie

Abstract Educational evaluation is an important and complex activity in the field of education. Rooted in the complicated, significant and multi-factor study of educational evaluation, the variables affecting educational evaluation can be summarized as two constraints: scientificness and publicity. The former, required to briefly grasp the key points of evaluation, is limited by value, space and time; the latter, required when focusing on educational equity, takes educational evaluation as public policy; it describes and analyzes the ways to realize publicity from the perspectives of participation and decentralization. The college entrance examination system in China, a sort of educational evaluation covering a vast range of aspects and involving a large number of stakeholders and the most complicated variables, is also constrained by both scientificness and publicity.

Keywords Educational evaluation · Scientificness · Publicity

Educational evaluation refers to:

> The systematic investigation into the process or outcomes of the implementation of a particular educational program, also synonymous with 'program evaluation': such investigations answer calls for accountability, assist in decision making, aid program development and planning, and serve the needs of research. Current approaches to evaluation (in this programmatic sense) stress a comprehensive, naturalistic methodology that goes beyond sole reliance on quantitative analysis (William 2003: 130).

Evaluation is an important and complex procedure in educational activities. It is important because it plays a guiding role in educational activities; it is complex because it involves various boundary conditions and variables, which are obscurely correlated. More importantly, as a value-assessing activity, evaluation is highly subjective. Evaluation research can be defined as the systematic assessment or investigation of the worth, merit or value of an innovation, an initiative, a policy or a program. As well as gauging the worth or value of innovation, evaluations have been used to measure the 'efficacy', 'effectiveness', 'efficiency' or 'impact' of interventions or initiatives (McCulloch and Crook 2008). Based on its significance, multiple

W. Xie (✉)
Distinguished Professor of Liberal Arts At, Tsinghua University, Beijing 100084, China

Y. Deng and B. Gao (eds.), *Educational Research in China*,
https://doi.org/10.1007/978-981-16-1520-7_16

variables and subjectivity, in reference to the classification of evaluation in *The Routledge International Encyclopedia of Education*,[1] variables and boundary conditions, which affect educational evaluation, are summarized as double constraints, namely, the constraints of scientificness and publicity. These constraints are requirements of evaluation, for efficiency and equity. They are complementary and consistent. Lastly, the evaluation reform of the college entrance examination (CEE) in China is taken as a case to display the working mechanism of the double constraints and their mutually restrictive relations.

1 The Scientificness of Educational Evaluation

The scientificness of educational evaluation implies a process that can assess how much an educational activity meets the needs of its subject or whether it has realized or potential values, as required by science. It has two basic implications. Firstly, considering the characteristics and requirements of science itself, this sort of educational evaluation has to be classified according to disciplines, that is, evaluating according to the types of educational activities. This is both a basic requirement of scientificness and a fundamental characteristic of science. Classification is a critical basic standard to ensure fairness in educational evaluation. Clearly, it would be inappropriate to apply the laws and standards of natural science education when assessing developments in humanities and social science education. Therefore, a reasonable classification of the evaluated objects is a prerequisite for a scientific-educational evaluation. Secondly, scientificness in educational evaluation also requires highly correlated variables or important variables of priority to be identified using approaches such as clustering from the various variables that affect educational evaluation, and thus display the values and functions of the evaluated objects. It must be pointed out that highly correlated variables should meet two conditions. First, they should be in a minor amount and not endogenous. For instance, when evaluating the educational positioning of a high school, a highly correlated variable would be the percentage of high school graduates admitted to college. If most high school graduates are admitted, it is a preparatory high school; if only a minority are admitted, it is a terminal high school (Trow 1961). Whether it is longed for by international youth is also a fundamental metric for measuring the level and quality of internationalization in a university. Too many variables or factors usually reveal an insufficient understanding of the core values of the evaluated objects. Second, there needs to be conciseness and immunity from ambiguity. For example, when assessing the teaching quality of a university professor from an academic perspective, a scientific variable would be an audit by peer teachers of the same discipline. Or, when assessing the developmental stage of higher education, the most recognized and convenient indicator is Gross

[1] The Routledge International Encyclopedia of Education classifies evaluation as two types. One is democratic, or participative, as opposed to non-democratic; the other is scientific, or experimental, as opposed to naturalistic.

Enrollment Ratio; 15% of GER meaning massification stage, 50% meaning universalization stage. Finding highly correlated variables like these is critical for ensuring the fairness in educational evaluation.

Generally speaking, the scientificness constraint of educational evaluation involves three significant restrictive factors.

Firstly, one must understand the values of the evaluated objects, that is, the most essential educational values of the object itself.[2] This is the most fundamentally restrictive factor in regard to ensuring the sensibility and scientificness of educational evaluation. As various educational variables are obscurely correlated, some scholars even believe that, "school is an unopened black box with unknown features…The variables and boundary conditions are complex. Their relationship is not a simply linear one" (Hallinan 2004: 50). Therefore, what matters more in the educational quality of schools and the academic development of students has always been an important question to answer in educational evaluation. Meanwhile, as required by science, different educational activities carry different values. Thoroughly and accurately revealing their core values through educational evaluation is the most important condition for the scientificness of educational evaluation. For example, to evaluate the research outcome, the number of published papers should not be the sole factor. Instead, the standard of representative publications recognized by academia should be employed, because representative publications are highly correlated with scholars' professional depth and abilities, while simple quantitative metrics are easily biased. Another example, evaluation in physical education is typical in identifying the core values of educational activities. Physical performance is not synonymous with physical fitness or athletic performance. It would be unfair for students in different physical conditions and it is not consistent with the core values of physical education itself. Mr. Liang Shuming[3] (2012) has a classic statement for this,

> Physical education is not training for military or other purposes. Neither is it merely for improving physical stamina and dexterity. It is for the improvement of the whole person. It combines education with exercises to develop what distinguishes between humans and animals, which entails not only physical growth because the stamina and dexterity human acquired from long periods of training will never compete with the stamina of a cow or the dexterity of an ape. However, humans transcend animals for the ability to harness their body using their will.

The training of will and the cultivation of self-discipline therefor are the exact essential purpose of physical education and are naturally regarded as the main criteria for the evaluation in physical education.

Secondly, the space factor in educational evaluation is important, that is, the restrictive effects that space characteristics of evaluated objects have on educational

[2]Two nouns are used in educational reference books when it comes to educational evaluation, namely, evaluation and assessment. This paper distinguishes educational measurement from educational evaluation or assessment and sees it more like a basic step of evaluation. While educational evaluation or assessment is a kind of valuation based on the results of measurement. This paper mainly deals with educational evaluation or assessment.

[3]Liang is a famous philosopher, educator, and leader in the Rural Reconstruction Movement in the late Qing dynasty and early Republicaneras of Chinese history.

evaluation. Evidently, the significance and values of evaluated objects are closely related to space they are in. It can even be held that different space characteristics have different education activity requirements. For example, in the realities of Chinese higher education, higher education institutions directly affiliated to the Ministry of Education (MOE) have different characteristics than those in local areas; this means that they have different objects of service and, thus, evaluation criteria. Therefore, evaluation criteria for the former cannot be replicated for the latter. During the educational evaluation of higher education institutions in the late 1990s, it was heatedly debated that how to assess the state universities in contrast to the provincial HEIs. Namely, whether the standards for "excellence" should be differentiated for the two different types of HEIs. After discussion, the *First MOE Expert Evaluation Committee on University and College Education* agreed that they should be the same. Despite they are differences in space characteristics and objects of service, being able to provide quality teaching services for their subjects should suffice for excellence. Educational level implies differences in service objects, while educational quality indicates the differences in service quality. Therefore, a top-tiered HEI can be of low quality because of its poor service; a low-tiered HEI can be of high quality because of the quality services it provides (Xie 2014). These differences are determined by the space characteristics of educational activities. An irrational evaluation would overlook the space characteristics of educational activities or the evaluated objects. More importantly, the educational evaluation that takes into account the space characteristics of educational activities could facilitate higher education institutions playing their role in guiding local economic and cultural development and impose higher standards on those under direct MOE affiliation.

Thirdly, there is a time factor in educational evaluation, that is, evaluating based on the time or time efficiency characteristics of educational activities. As it is known, time or time efficiency is a remarkable characteristic of educational activities. Entrusted with expectations from society, it usually takes a long time for education to fully register its impact. However, the time or time efficiency characteristics of educational activities register differently with different majors or disciplines and therefore it becomes a confining factor for evaluation. For example, the research findings of science and engineering disciplines and those of humanities and social sciences disciplines are strikingly different in terms of time or time efficiency. Similarly, compared to the evaluation of applied research, the evaluation of fundamental research takes longer, often more than a decade or even decades. One research finding, in particular, explains very well the difference caused by disciplines. Robert King Merton, a prestigious American sociologist, introduces a comparative study on the citing of papers of different disciplines. The citations from natural science papers are found to be mostly from papers published within five years from their time of publication. For example, in physical science, as represented by *Physical Review and The Astrophysical Journal*, 60 ~ 70% of citations are from publications of the recent five years, while only a few are from publications from five years ago or even earlier. As for social sciences, citations are seldom from publications within the last five years. For example, as shown in *The American Sociological Review*, *American Journal of*

Sociology and *British Journal of Psychology*, 30–50% citations are from publications from the last 5 years. However, in humanities and liberal arts, citations from publications from less than 5 years ago account for the least percentage. For instance, in *The American Historical Review*, *Art Forum* and *The Journal of Aesthetics and Art Criticism*, 10–20% of the citations are from publications within the last 5 years, whilst most of the citations are from publications produced 5 years ago or even earlier (Merton 2006). This research finding is strong proof that there is a great difference in the time cycle between the social sciences and the humanities and liberal arts. The former tends to be shorter, the latter longer. They cannot be treated equally. Thus, the time or time efficiency characteristics of different objects should be considered, and those of different disciplines should be respected. Since the amount of citations now weighs more in educational evaluation, this finding serves as a great reference for improving our mechanism of evaluation.

Apparently, scientificness is a basic constraint of educational evaluation and the basis for ensuring its reasonableness. To achieve scientificness, the significance and values of evaluated objects should be accurately understood, scientifically categorized and well summarized.

2 The Publicity of Educational Evaluation

The publicity of educational evaluation refers to the attributes and characteristics common to public policies that educational evaluation has. In other words, in the category of public policies, the educational evaluation must be subject to their constraints and regulations. Determined by the spillover effects (or externalities) of educational activities, the publicity of educational evaluation is closely related to the fundamental attributes of educational activities. Well-known economists like Milton Freidman and Robert Lucas have explained and discussed the spillover effects of educational activities from the perspective of their producers and recipients (Xu 2015). Domestic scholars have also distinctively pointed out that with its considerable external benefits, education can benefit both individuals and society greatly, either economically or otherwise (Xu 2015). There are two types of spillover effects of education. It can be remarkably positive, or negative because of its intrinsic problems. For this reason, the evaluation of educational activities cannot afford to be overlooked as a critical subject matter in public policies. Especially, in modern society where education is universal and critical to people's development, it has become a matter of interest that draws direct and realistic attention (Xi 2013). In many countries, education usually receives special attention from country leaders and is very attended to in national public policies. It has even become a significant discourse given by state leaders and their administrations.

The scientificness of educational evaluation requires educational activities to be efficient, while its publicity requires them to be fair. Moreover, they are unified. In modern society, educational equity is an essential part of educational quality

and a basic requirement for achieving reason in educational activities. An important goal of educational modernization is ensuring educational equity and continuously improving its quality. Therefore, educational quality is an essential part of the publicity of educational evaluation. One of the main goals of educational activities, and their publicity, is to maintain and promote educational equity and thus improve the scientificness of educational evaluation more effectively. The progress of time and education in and of itself have furthered the features and the implications of educational equity. In the twenty-first century, the value of the publicity of educational evaluation has at least three new layers of meaning for educational development.

Firstly, educational equity has become an independent variable that affects individual development and social progress. The conventional theory and idea of "basic equity," the basis of educational equity, is undergoing remarkable changes. According to H. Beck, the so-called "basic equality" is the principle which is fundamental to all egalitarian theory. "It is the idea that all human beings are equal in dignity and worth, and are therefore equally worthy of concern and respect" (Hallinan 2004: 119). Equity here is mentioned as the effect in the cause-and-effect relationship. Therefore, "it defines equity in fundamentally negative terms therefore, in terms of freedom from, rather than freedom to" (Hallinan 2004: 119). Educational equity is now defined in more positive terms, namely, it is no longer the effect in a cause-and-effect relationship or a "dependent variable," but an "independent variable" in the development of mankind. This demonstrates the ever-greater emphasis and attention that educational equity now receives. Among the five areas concerned with people's wellbeing, education, health care, employment, income distribution and social security in our society, education has become a primary issue (China Daily 2013).

Secondly, the intrinsic differences of education itself have posed great challenges to the publicity of educational evaluation. Undoubtedly, the 9-year compulsory education and the massification and upcoming universalization of higher education mean a remarkable step forward for educational equity. Education is now far more accessible. However, improvement in the quality and extent of educational participation presents more serious challenges to the public nature of educational evaluation. Because the improvement has enlarged education's intrinsic differences, including the differentiation of education itself, the different demands of different stakeholders, the fundamental differences in educatees' academic and literate level, etc., how to better meet the demands of different stakeholders and educatees with different personalities and in different conditions and how to promote the common development of people and social equity will become the new task for educational equity. Especially during a social transition, people would pay more attention to educational equity and have greater demands for it since they would observe the group as a whole and through comparison, have an intensified sense of deprivation.

Thirdly, equality of outcome is becoming the main goal in the pursuit of publicity in educational evaluation. Amid the social development and educational reforms in the twenty-first century, the goals of educational equity are changing. According to some researchers, equality of opportunity is determined by the different levels of education received by the less advantaged groups of people in a stratified society and educational system. Equality of participation, a stage more advanced than equality of

opportunity, is measured by participation rate rather than access, and/or the quality and nature of experience and participation in different educational departments and areas by different groups of people, whilst equality of outcome or success is measured by the educational success or the educational success rate of marginalized groups. As participation in certain fields of education, notably university education, is often defined as the pinnacle of educational success—implicitly if not explicitly (Hallinan 2004: 121). Clearly, China now faces challenges with its CEE reforms, it is proof that this sort of equality of outcome is becoming the goal of educational equity. Therefore, the publicity constraint of educational evaluation must coordinate with the demands of different stakeholders in education, continuously strive for equality of outcome whilst ensuring equality of opportunity and participation, so it can achieve as much educational equity as possible in the society. The publicity constraint of educational evaluation plays a guiding role in educational reforms, facilitates the equitable distribution of education resources, governs the behavior of various stakeholders and resolves conflicts, thus promoting the development of education. Not a necessarily appropriate analogy: the publicity of educational equity is like the dishes of a feast, recognized by guests with different tastes. The key is not necessary to ensure the quality; instead, it is more important to discover the common ground shared by the guests, which entails a certain degree of cooperation and compromise. This happens to be one of the functions and values of public policies.

Generally speaking, the publicity of educational evaluation and the formulation of relevant public policies can be realized in two ways. First, through an open, fair and standard procedure where representative stakeholders participate; the subject matter of one educational public policy is thoroughly discussed in order to achieve as much consensus as possible. This kind of public engagement is an important channel of publicity of educational evaluation, a necessary procedure in ensuring and improving its reason and a fundamental social foundation for formulating its criteria. For instance, to formulate the evaluation standards for the titles of professor and associate professor, considering that various disciplines and academic fields are involved, the university first submits the drafted articles to some faculty in different schools and departments for discussion and guidance, revises them based on the general requirements of the university and guidance from various sides, and finally puts them to a vote by the entire faculty of related schools and departments. Then the evaluation guidelines and documents can be finalized after a majority consents, as required. In another instance, to accommodate geographical and disciplinary differences and reach a consensus on the standards, departments in charge of execution invite experts and representatives for discussion, consult and revision. Finally, the evaluation scheme is formed based on suggestions and guidance from various sides.

Second, to implement and showcase the publicity of educational evaluation through the delegation or division of powers. When the differences conflict and the divergence of externalities from certain educational activities are so significant that the consensus or a unified evaluation scheme is hard to achieve through proper procedures, delegation or division of powers are needed for the implementation of educational evaluation. Delegation or division of powers in educational evaluation

means that in order to maintain the reasonableness of educational evaluation, a delegation or authorization mechanism is employed by evaluating agencies at various levels and of different types to evaluate educational activities and educational institutions. This kind of delegation or division of power is necessary and justifiable in educational evaluation. Especially so when evaluating areas or departments of a strong professional nature; it is difficult to reach a consensus regarding the evaluation criteria and schemes for various professional educational activities, making it necessary to trust or authorize professional institutions to form an evaluation. For instance, to develop the disciplines of liberal arts which involve various disciplines and fields including management, economics, law, literature, history, philosophy, and art, a university specifies that on the general basis of "quality first," academic evaluation should be carried out by the basic academic unit so that the distinctive characteristics of each discipline should be respected. Instead of forming uniform evaluation criteria and requirements for all schools or departments of liberal arts, the university allows each department/school/institute to have its own criteria and requirements, achieving general diversity in evaluation (Xie 2011). It must be pointed out that despite certain differences, this sort of delegation or division of powers is more equitable and reasonable. The reason being, adapting to various disciplines or majors which have their own features and rules, builds more practical criteria, which can be more easily recognized and thus better promote educational reforms. There are no unified requirements for this kind of delegation or division of powers. Neither is it pure laissez-faire. Usually, it is a combination of centralization and decentralization. In the above case, the university proposes the requirements "regarding it at the national level, striving for impact worldwide, standing the test of history and contributing to practices," the guiding role of "three impacts," namely, academic impact, social impact and international impact, and emphasize the "intersection" between the development of liberal arts and the country, as well as between a nation strong in higher education and a world-class university (Xie 2011). Practices have shown that educational evaluation with this kind of oriented delegation or division of powers usually promises a better chance at reason and equality, and thus promotes educational reforms more effectively.

It can be concluded that the publicity and scientificness of educational evaluation are not contradictory. Instead, they are unified. Because publicity helps to maximize and balance shared values in educational evaluation and earn recognition as widely as possible. This happens to be what the scientificness of educational evaluation intrinsically requires.

3 The Double Constraints of CEE Reform

The CEE system of China is an educational evaluation activity covering the largest majority population and involving most stakeholders; it is influenced by the most complicated variables in the Chinese educational system. The double constraints of

educational evaluation are essential to the scientificness, equity and reasonableness of the CEE system.

3.1 The Scientificness Constraint of CEE Reform

The scientificness constraint of CEE reform mainly lies in the formulation of various policies including the creation of tests, scoring, admission and so forth. It is governed by the regular patterns of the system. Among them, the foundation of the system, are the core values of the examination itself.

Undoubtedly, as higher education evolves from the phase of massification to the phase of universalization, as the connection between basic education and higher education, the CEE system also plays a role in selecting talents and guiding education whilst fundamentally fulfilling its basic role in moral cultivation. These two are both fundamentally important functions of the system. However, how to bridge and coordinate them organically is one of the main contradictions within the system and the greatest challenge and difficulty it faces during test question design, test organization and systemization. The reasonableness of the examination basically relies upon whether these two basic functions can be scientifically coordinated. It has a direct bearing on the fulfillment of the fundamental functions and core values of the CEE system; it is also the most essential implication and object of its scientificness constraint.

Despite that basic education and higher education both fall in the category of education and follow common educational rules, they are different. They are at different levels and of different types and follow different principles, orientations and approaches. The reform of the content of CEE is an essential part of the system. It connects the curriculum systems, curricular materials and forms of organization of basic and higher education. The scientificness constraint of the CEE requires a scientific cohesion and coordination of the two.

From the perspective of the CEE content reform, scientificness has a role to play in the following three aspects.

Firstly, to connect and coordinate the curricular materials of basic and higher education and to seek "intersection" between the two in content. These two types of education are remarkably different in terms of the level of curricular materials. Basic education pays more attention to general knowledge, whilst higher education emphasizes professional knowledge. The former pays more attention to the cultivation of people, whilst the latter emphasizes the cultivation of talents. The former pays more attention to the acquisition and training of basic knowledge and abilities, whilst the latter emphasizes the investigation of advanced knowledge and the training of research abilities. The list goes on. Coordinating the two kinds of curricular materials and finding their intersection is an apparent and significant aspect of the reasonableness of the CEE content.

Secondly, one ought to connect and coordinate the curricular logic of basic and higher education, and to seek "intersection" in curricular orchestration. Clearly, the

curriculums of basic and higher education are orchestrated in different ways and follow different logics, which leads to differences in the teaching and acquisition methods and curricular evaluation. In regard to this, the renowned American educationalist John Dewey has an important statement. He points out that every study or subject has two aspects: one for the scientist as a scientist and the other for the teacher as a teacher. These two aspects are in no sense opposed or conflicting. But neither are they completely identical. For the scientist, the subject-matter represents a simple given body of truth to be employed in locating new problems, instituting new researches, and carrying them through to a verified outcome. To him the subject-matter of the science is self-contained. He refers various portions of it to one another; he connects new facts with it. He is not, as a scientist, called upon to travel outside its particular bounds; if he does, it is only to get more facts of the same general sort. The problem of the teacher is a different one. As a teacher, he is not concerned with adding new facts to the science he teaches, either in propounding new hypotheses or in verifying them. He is concerned with the subject-matter of the science as representing a given stage and phase of development of the experience. His problem lies in inducing a vital and personal experience. Hence, what concerns him as a teacher, is the ways in which that subject may become a part of the experience; what aspects there are in the child's present that may be useful to refer to; how such elements are to be used; how his own knowledge of the subject-matter may assist in interpreting the child's needs and doings, and determine the medium in which the child should be placed in order so that their growth may be properly directed. He is concerned, not with the subject-matter, but with the subject-matter as a relational factor in the overall growing experience. Thus, to see it is to psychologize it [16]. It can be seen that the coordination and "intersection" of the two curricular logics are intrinsically required for the reasonableness of the exam. To reform the CEE and improve its reasonableness, the intersection between the psychological logic of the curriculum of basic education and the scientific logic of the curriculum of higher education should be understood and identified, so that they may be coordinated in an organic manner.

Thirdly, one must connect and coordinate the curricular forms and structures of basic and higher education, and seek an "intersection" with the form and structure of the CEE content. It can be seen that basic and higher education is organized in different ways. Although their curriculums are both related to "disciplines" or "subjects" (in Chinese they are both written as xue ke), they mean different things. Subject in basic education equals to curriculum, and the evaluation of subject equals to curricular evaluation. While in higher education, the curriculum is closely related to many disciplinary factors, is an important aspect of the discipline. Curricular evaluation is part of the disciplinary evaluation. Subject and discipline, the different words of choice, exemplify their differences in form. Subject, often used in "subject matter," has a closer relation to knowledge. While with the connotation of controlling one's behavior, discipline stresses the boundary conditions and limits of knowledge. This difference is significant and embodies the important features of these two different kinds of education. It reflects different learning modes and different educational patterns at different stages of physical and psychological development. Certainly,

the CEE reform and its reasonableness also require it to coordinate the two forms of knowledge, find and locate their "intersection." To put it more straightforwardly, this kind of coordination and "intersection" should be demonstrated by the structure of test papers and the type of questions it presents. It is only by coordinating and merging the knowledge which has two different logics and forms, relevant learning methods and the ways of thinking that the CEE mechanism displays in its role in talent recruitment and teaching guidance, can it realize its core values.

3.2 The Publicity Constraint of CEE

The publicity constraint of CEE mainly refers to the procedural constraint on the CEE equity, including the choice of tested subjects, the proportions of compulsory and elective subjects, the policies of CEE and the independent recruitment exam of many universities, the weighting of weighted scores, the time scheduling and other procedural requirements for the fairness of the relevant rules and regulations. The CEE system concerns the interests of thousands of families. It is a highly sensitive evaluation activity in terms of educational equity. To some extent, it even matters to national and social stability. Therefore, the publicity constraint is a key process and requirement for the reasonableness of the CEE system.

In fact, due to the imbalance of social and educational development and the division of interests, despite being restricted by the publicity of educational evaluation, the CEE reforms and improvements cannot be unanimously recognized by different social classes, geographical regions and stakeholders across society. CEE reforms in some provinces are possibly more criticized for public reasons than for scientific reasons, which makes the constraint of publicity even more prominent. The CEE system is one of the most important public policies in a country and a society. The ability to earn as much common ground as possible across different social classes, regions and stakeholders across society has a direct bearing on its success and reasonableness. Taking into consideration this orientation, requirement and the current realities, the publicity constraint of the CEE system should be implemented in the following aspects.

The first aspect should be public engagement. Public engagement is one of the main forms and approaches through which the system administers its constraint of publicity. One of the basic practices is to invite representatives from different groups or stakeholders to a joint exchange and consult on CEE reform and to solicit opinions from various groups, including those with different stances. In this process, different viewpoints can be exchanged and discussed. The goal is to coordinate the appeals of different stakeholders and reach a broad consensus. The forms and channels of public engagement can certainly be diverse. A typical example would be a social evaluation of two educational loan programs of the World Bank. To push the educational loan programs to not only improve educational efficiency but also achieve equality to the greatest extent possible and expand the beneficiary coverage, the World Bank authorized an independent third-party panel to execute a social

evaluation on the feasibility of studies of involved schools and institutions. To put it simply, in this basic form of social evaluation, the stakeholders of related projects are scientifically sampled, then the opinions of various parties, mainly their views and understanding on the pros and cons that the project has on each of them, are solicited through questionnaires, interviews and other forms of fieldwork. Then, the views are summarized and concluded to devise a feasibility report for relevant schools and departments. Constructive revisionary advice is also proposed. Moreover, this kind of social evaluation is an important reference for the World Bank Executive Committee in its decision-making process.

Secondly, there should be a delegation in the CEE system, including delegation across areas, types, levels and so on. The goal of the publicity constraint is to reach consensus across different areas, types and levels; the CEE system has to not only satisfy nationally uniform requirements but also accommodate the different characteristics of different areas. This is another contradiction within the system and the most vital aspect of its publicity constraint during procedure and regulation formulation. However, as society further divides, educational differences grow, people begin to have the diversified pursuit of interests. Moreover, different areas are imbalanced in their level of basic education development. Therefore, it is hard to achieve general or major consensus within the CEE system and the exam content through simple public engagement. In this circumstance, the publicity constraint has to be exercised in the form of delegation, making it the second basic approach.

There are four basic types of delegation in the CEE system. First, delegation across regions. Namely, for some reason, certain provinces can hold independent entrance exams in a manner of limited delegation or authorization, just as those holding a few years earlier in certain provinces, municipalities and autonomous regions. Second, creating test papers independently. For particular reasons, some places cannot be included in the national exam and need guidance and assistance from the government. They can design their test papers independently from the national exam with the help of relevant government departments and thus accommodate the demands of local educational development and realize the fundamental functions and core values of the CEE system. For instance, several sets of tests have been created over recent years in this way to meet differentiated demands. Second, the delegation is across different levels. The CEE system is delegated across the different levels within the higher education system, that is, institutions of higher education directly supervised by the MOE, local universities, local vocational and technical colleges. For instance, institutions of higher education directly supervised by the MOE recruit nationally and serve the nation as a whole, and therefore have to take the national-level exam. Certain key disciplines or majors of local universities, having met required conditions, can voluntarily apply to join the national level through certain procedures. While local institutions of higher education mainly recruit and serve at the local level, which requires them to accommodate local characteristics. Therefore, they are not required to join the exam system at the national level and can recruit autonomously through appropriate delegation or authorization practices. While some vocational or technical colleges can even employ the registration enrollment scheme provided that fairness and quality are ensured. Lastly, the CEE for higher vocational education can,

according to the characteristics and demands of vocational education, employ delegation, authorization or independent test development and other appropriate forms to better accommodate its requirements for development. This sort of delegation is by no means entirely fixed. It can be a form of delegation or authorization that is fused with the NCEE system or the CEE content.

It has to be pointed out that delegating the CEE system or its content is not forsaking educational equity. It is done precisely for the realization of more reasonable educational equity. One, it is difficult to achieve educational equity by applying uniform content to different regions, levels and types during evaluation. Moreover, from the theoretical perspective of educational equity, delegating the CEE is more in line with how different stakeholders in modern society subjectively perceive educational equity. This is because according to the basic theories of educational socialism, when society divides and regional development is biased, in concern of the pursuit of educational equity, people prioritize equity between groups over equity within groups. Theories of educational equity also uphold this order of priority, as people tend to refer to their surroundings when judging and evaluating educational equity.

Of course, this kind of delegation is not laissez-faire. It is governed by uniform guidance and principles; it is decentralization combined with centralization and delegation. The governmental departments in charge have the right of direct guidance and supervision over the decentralization of the CEE system, which is also their responsibility.

Thirdly, there should be stability within the CEE system; this is an important aspect of the publicity constraint of educational evaluation or the CEE. This requirement of stability aligns with the basic characteristics and rules of educational evaluation and meets people's demands on the reform and betterment of the CEE system.

The CEE stability can be interpreted in at least two aspects: firstly, reform on the CEE system and content cannot be frequent, let alone changeable. There should be a bottom line for its cycle or a time limit. As stated by Laozi in *Tao-Te Ching*, "Running a large country is like cooking a small fish." In a country as populous and imbalanced developed as China, stability deserves even more attention. As validated by theories and practices, unduly frequent CEE reforms could only bring negative influence on educational development. Secondly, consistency is necessary for the reforms on the CEE system or its content, which are required to keep up with social development, scientific and cultural progress and educational reforms. After all, the CEE in the early twenty-first century has a drastically different historical background and social pursuit than it did in the late 1970s. But consistency must remain within the CEE system and its content reforms. The changes which are made should mainly be minor, major ones are to be avoided. There is no doubt that China's CEE reforms should learn from the progressive mode that economic and social reforms undertook. In the meantime, it should avoid points occurring periodically over the accumulation of contradictions and problems as well as the "stresses" that explode at these points, which occurred in the progressive mode of economic and social reforms. This is what distinguishes educational reforms from the economic and social ones. Therefore, CEE reforms including those on its content should be progressive and gradually, in

a well-planned manner, eliminate various conflicts which have accumulated along the process.

Essential to the publicity constraint, stable CEE reforms aim to provide relatively stable social expectations for the public and the stakeholders, as well as a relatively stable institutional environment for the reform and development of China's basic and higher education. This is a foundation for a well-designed CEE system.

Clearly, the publicity and scientificness of the CEE evaluation system are consistent. This is especially so, as both the public engagement in CEE reforms and the delegation and stability benefit the CEE evaluation system in terms of efficiency and scientificness. Imagine a CEE evaluation system that fully reflects the different appeals of basic and higher education, coordinates the tension between different variables, and embodies the characteristics of schools and regions at different levels and of different types as well as the inherent requirements for stable educational activities. Certainly, it can better coordinate the purposes of talent recruitment and teaching guidance, adapt to the demands of institutions of higher education and regions of different levels and types, enhance the degree of recognition and satisfaction from the public and the various stakeholders with the CEE evaluation system and the maximization of its values.

Educational evaluation is an extremely complex educational activity that involves a manifold of variables. The CEE reform is a public policy that receives the widest attention and can produce domino effects. Its constraints of scientificness and publicity raised in this paper, are fundamentally vital.

References

Anthology of John Dewey on Education. (2013). Beijing: People's Education Press. (in Chinese).

Hallinan, M. T. (Eds.). (2004). *Handbook of the sociology of education.* Shanghai: East China Normal University Press. (in Chinese).

Liang, S. (2012). *Education and life: Anthology of Liang Shuming on education.* Beijing: Contemporary China Publishing House (in Chinese).

McCulloch, G., & Crook, D. (Eds.). (2008). *The routledge international encyclopedia of education.* London: Routledge Taylor & Francis Group.

Merton, R. (2006). *Social theory and social structure.* Nanjing: Yilin Press (in Chinese).

People's Daily. (2013). The third plenary session of the 18th CPC Central Committee Kicks off in Beijing. *People's Daily.* (in Chinese). Retrieved October 9, 2020, from https://politics.people.com.cn/n/2013/1113/c1024-23520857.html.

Trow, M. (1961). The second transformation of american secondary education. *International Journal of Comparative Sociology, 2*(2), 144–166.

William, J. (2003). *The greenwood dictionary of education.* Santa Barbara, Calif: Greenwood.

Xi, J. P. (2013). CPC Central Committee Announces Decisions on Major Issues Concerning Comprehensively Deepening Reform. Xinhuanet. (in Chinese). Retrieved October 10, 2020, from https://www.xinhuanet.com//politics/2013-11/15/c_118164294.htm.

Xie, W. (2011). Three firsts and policy guidance. *Guangming Daily.* (in Chinese). Retrieved October 9, 2020, from https://epaper.gmw.cn/gmrb/html/2011-11/25/nw.D110000gmrb_20111125_5-14.htm?div=-1.

Xie, W. (2014). Is education level equivalent to education quality? *China Education Daily*. (in Chinese). Retrieved October 9, 2020, from https://theory.rmlt.com.cn/2014/1122/347729.shtml.
Xu, C. (2015). Spillover Effect of educational investment and its internalization. *Education Research Monthly., 3,* 40–47 (in Chinese).