ISSN 1346-6151

Japanese Journal of Evaluation Studies

Vol. 18, No. 2, June 2018

English Version

CONTENTS

Japan Evaluation Society

Office
The Japan Evaluation Society
c/o International Development Center of Japan,
Shinagawa Crystal Square 12th Floor, 1-6-41, Konan,
Minato-ku, Tokyo, 108-0075, Japan
E-mail: jes.info@evaluationjp.org

Japanese Journal of Evaluation Studies

Vol. 18, No. 2, June 2018
English Version

CONTENTS

【Foreword】

Semi-Special Issue : School Evaluation Studies from International Perspectives

Akihiko Hashimoto
National Institute for Educational Policy Research, Japan

This semi-special issue consists of four papers from Australia, Korea, Taiwan, and Japan, all dealing with the school evaluation system. Each of them is based on papers presented at the International Session of "Whole School Evaluation: Approaches used in School Systems in Australia, Korea and Taiwan" at the Japan Evaluation Society's 17th Annual Conference at Hiroshima University on November 26th 2016.

The term "Whole School Evaluation" (WSE hereafter) is quite new to educational research in Asian countries but has been a familiar term in certain countries since the 1990s. It seems to have started as a way to make a school accountable a whole but has shifted to place more emphasis on a school's self-improvement strategy. At the 2016 international session mentioned above, the notion of WSE was introduced in a report by Dr. John Owen explaining the case in the State of Victoria in Australia. In contrast with the international efforts of WSE, reports from Korea (by Dr. Sung Jae Park), Taiwan (by Dr. Shu-Huei Cheng), and Japan simply presented their own short history and the present circumstances surrounding school evaluation in their own country. The three reports from Eastern Asia all explained how their school evaluation process focuses on each school as a 'whole' and their system tends to emphasize self-evaluation rather than external-evaluation. However, it was hard to determine if East Asian schools were using a type of WSE or not.

Papers in this issue basically have the same content as the papers presented at the session in 2016. I must apologize because the editing process took time and the publishing has been delayed due to my velocity of work. Although time has passed, the international comparison of WSE and the East Asian school evaluation system remains unaccomplished. However, we have a precious opportunity to examine these four systems of school evaluation, and the task of comparing them from the viewpoint of WSE is an attractive and important theme to pursue.

In many countries, regions, or states in East Asia, "there could be further questions on the effectiveness and efficiency of the current evaluation system" (expression borrowed from Dr. Park's paper). This seems to be a "feeling" experienced in many schools in East Asia, as Dr. Cheng might agree as she states in her paper that schools in Taiwan have the task to "overcome any challenges that hinder effective self-evaluation", which sounds familiar to Japanese schools. This "feeling" must be studied scientifically. However, I hope this semi-special issue will provide a starting point for new research studies investigating methods of school evaluation.

【Article】

Whole School Evaluation: Approaches Used in the Public School System in Victoria, Australia

John Owen
Centre for Program Evaluation The University of Melbourne
j.owen@unimelb.edu.au

Abstract

While there is ample literature regarding theory underlying whole school evaluation (WSE), there are few examples of implementation. This paper provides a case example in one jurisdiction; the public school system in the State of Victoria, Australia. The paper shows how a commitment to mutual accountability underlies implementation success. On one hand, the State Department of Education (DoE) encourages individual schools to use WSE for improvement purposes. On the other, each school provides outcome information that enables the DoE to aggregate outcome information across the system, for accountability purposes. The paper outlines key strategies that have led to the success of this systemic arrangement.

Keywords

Whole School Evaluation, Educational Systems, Mutual Accountability, Improvement

1. Introduction

In all evaluation work it is important to identify the thing or object that is being evaluated. Perhaps the most popular evaluation object is a program, it is certain that much evaluation theory assumes programmatic principles, such as a set of objectives, the translation of these objectives into action and resulting improvements for program participants. However, in principle these assumptions can also be applied to the evaluation of an organisation and this is the basis for planning whole school evaluation practice (WSE).

2. Whole School Evaluation (WSE)

WSE can be defined as the use of systematic investigation of the quality of a school and how well it serves the needs of its community (Sanders and Davidson 2003).

Those responsible for WSE need to develop valid tools that acknowledge the added complexity of evaluating an object that is more complex than a program. Those involved in evaluation studies must first agree on the purpose of such studies. Most would agree that WSE would aim to improve the both the delivery of structures and processes, such as teaching, and student outcomes.

While a goal of WSE is to improve individual school quality, there is often a second purpose, that is to enable a school to be accountable to a higher authority in a system, such as a central education department and policy makers. In general terms, accountability is predicated on the assumption that government and citizens have the right to know whether programs funded from the public purse are making a difference.

3. WSE in the State of Victoria, Australia

3.1 Contribution to School Operational Reform and Strategic Planning

About 20 years ago, in the State of Victoria, a recently elected conservative government set up a major program of structural, curriculum and accountability reform which was designed to change the ways schools operated. While there was a wide range of reasons for this, a major impetus was a report by the Victorian Commission of Audit that reported that there was virtually no systematic data on the performance of the government school system in the State. It recommended to the Government that the education authorities should arrange for periodic, independent reviews of school performance against a wide variety of indicators and standards.

There have been a series of policy frameworks related to whole school evaluation over the past two decades. I will now outline the characteristics of the present framework, and compare them with key conceptual principles.

As of 2016, in the State of Victoria, one can think of a school in the government system being involved in a Strategic Planning exercise: an evaluation/ development cycle involving the following stages:

Stage 1: School Review
Stage 2: Setting and Prioritising Goals
Stage 3: Development and Planning
Stage 4: Implementation, and
Stage 5: Monitoring (self-evaluation)

The School Review (Stage 1) is a periodic assessment of the performance of the school that allows the system to be satisfied that the school is meeting key State level objectives. These are in the areas of:

1) Student achievement
2) Student engagement and
3) Student wellbeing.

Each school is also encouraged to develop additional initiatives that respond to particular needs.

In addition to providing evidence of existing performance, the findings of the School Review are designed to set the direction for operation of a school over its next four years of operation.

Stages 2 and 3, involve the development of a document, now known as a Strategic Plan. The Plan is designed to link planning goals to implementation strategies (Stage 4).

Stage 5 can be thought of as a complement to the School Review and is known in the DoE documents as self-evaluation.

3.2. Key Evaluative Components of WSE: Stages 1 and 5. School Review and Monitoring by Each School

Stage 1. School Reviews provide an analysis of current school performance and practice, and should make a positive contribution to the school's efforts to improve student outcomes. A Review is likely to suggest teaching and other educational strategies that might be changed or introduced in order to improve the levels of achievement across the school. Reviews involve an accredited peer reviewer and the involvement of two principals from another school. This group could be thought of as constituting an Expert Panel who may consult staff, students and community members.

Peer learning is an objective of the strategy, providing an opportunity for the school leadership to build their knowledge and capacity, and to learn about other successful school practices that improve student achievement, engagement and well-being. In addition, a senior staff member known as the Senior Education Improvement Leader (SEIL) is expected to coordinate the Review process, and engage with the school's previous self-evaluation findings and other relevant data. Members of the School Council are expected to endorse the terms of reference of the Review, be involved in the processes, and receive and endorse the Review findings. They are also expected to provide feedback to the Department of Education on the quality of the Review process.

Stage 5. Monitoring via Self-Evaluation involves assembling evidence from community consultations and analysis of student outcomes over a yearly period. A school is expected to draft an annual report of performance and present it to the school community: staff and parents. Again, the school leadership and in particular the Principal is expected to manage the process with the support of the SEIL. On the basis of what we know about the resources required to manage evaluations, the role of the SEIL is crucial.

According to a DoE document on Strategic Planning the SEIL is responsible for

1) supporting the school with planning self-evaluation
2) providing support to analyse and interpret data and challenge where necessary
3) provide input into and feedback on the outcomes of each self-evaluation, and
4) endorse outcomes of the self-evaluation.

3.3 Data Usage by DoE

The evaluative approach adopted in the School Review largely separates out cause and effect or outcomes and processes. This is an approach that has its genesis in the work of a key theorist Joseph Wholey. In summary, outcome data is provided to schools or analysed on their behalf, by the DoE. By recourse to the the professional knowledge of educators, information is assembled about educational processes that are implemented by the school. It is also expected that key members of the Review and school staff will suggest new or additional processes that could be implemented in the future (the next round of Planning/Implementation) that could improve student outcomes.

This approach is only possible because the DoE takes the major responsibility for the collection and analysis of outcome data on a systematic basis across all schools, provides feedback to each school about its performance, and prepares comparisons of this performance with 'like schools' those with similar SES (socio-economic status) levels.

3.4 Range of Outcome Information

Each school receives from the DoE or assembles the following outcome information.

1) Enrolments by year level over time
2) National achievement outcomes in English/mathematics/science skills over time based on teacher assessments based on national curriculum objectives, and a National testing program (NAPLAN).
3) Student Attitudes to School
4) Student Absences Rates and Causes

5) Development Status of Children Entering School (Prep Entry)
6) Social, Emotional and Behavioural Data on Children Entering School (Prep)
7) a School Performance Summary, including analyses of student outcomes and in addition,
    •Overall SES Profile
    •Proportion of ESL students
    •Results of Parent Satisfaction Survey
    •Results of School Staff Survey

As indicated this information allows school leadership to compare their performance with 'like schools'.

4. School Usage of Performance Information: Role of the Principal in Effective WSE

It is clear that the school Principal is a key to making the evaluation/planning cycle work, and its influence on school decision-making (Ikin and McClenaghan 2015). Principals need to be data literate, with knowledge about the conventions of interpreting and using findings. Being a school leader now requires a positive attitude to accepting evaluative information, and to create a culture of staff and organisational learning (Earl and Katz 2001). Principals also need to balance these needs with the ongoing day-to-day decisions that relate to the operation of the school as a whole in terms of the delivery of the Strategic Plan.

Can this work in practice? The following case study provides some answers to this question.

Aix College is a year 7-12 school for girls located in Melbourne. The College is one of approximately 400 secondary schools in the government education system in Victoria, which falls under the Department of Education (DoE).

For about 20 years the Victorian education system has encouraged considerable devolution of authority to schools, and to school leadership. While there is considerable autonomy, all schools are required to develop a four-year Strategic Plan and to collect information that will be reported to the Department.

The Strategic Plan summarises how curriculum will be delivered over a period of four years, taking note of more general system level guidelines. Schools are expected to monitor the Plan during this time, making adjustments as they see fit. Towards the end of the period, a mandated evaluation is undertaken designed to assess the success of the Plan, and provide information that can be used for the next planning period.

A range of information must be collected during the period of a Plan. Data which must be to be reported to DECS includes annual surveys of parents, staff and students. DECS analyses these data across across all schools, providing an overview of the school education system over time. National tests of student literacy and numeracy are also administered. Thus, DECS has a database for all schools on these measures as well as others, such as student attendance and year 12 student performance.

Results are also provided to individual schools. This allows Aix College to compare its performance to schools in the same SES band.

Thus the performance management process serves both accountability and self-improvement purposes.

Aix College is led by an experienced leadership team; a Principal and two deputies. The Principal has a one-line budget, of several million dollars. While a majority of funding is provided by the State, additional funds are raised, some by a range of entrepreneurial activities, which are encouraged by the Department of Education. A major expenditure relates to hiring of staff over which the Principal has control.

The school operates at three levels: the school as a whole; departments made up of teachers working in the same pedagogical area (eg science); and at the individual teacher level in classrooms. A key role for the leadership team is to remind teachers about the Strategic Plan, which summarises school priorities for each four-year period.

The Principal reports to the School Council which signs off on the Strategic Plan. In addition to information about its progress, the Principal reports summary findings from surveys and student performance on a regular basis. There is considerable interest among Councillors about the performance of the school. An important role of the Principal is to interpret data patterns and to report decisions based on these patterns. Of particular interest is the progress of key strategies in the Strategic Plan; an example was a 'numeracy across the curriculum' initiative. Additional normed tests were used by teachers to check on student progress as a basis for assessing the effectiveness of new classroom strategies.

Key issues for the school leadership are student outcomes and quality of teaching. In the past the Principal was of the opinion that teachers were 'accountable to themselves'. Now all teachers undergo an annual review that result in individual performance plans that are signed off by the Principal, teachers are encouraged to become involved in analysing data that is relevant to their own teaching. This has had mixed success with some teachers resisting the use of data. So, the leadership has employed an external consultant to work with teachers on meaningful improvement strategies. This includes the use of a formative survey designed to indicate how a teacher can adopt new approaches. These procedures are confidential between the consultant and each staff member.

The Principal acknowledges that the use of data in making management decisions involves a lot more work than was the case before schools were required to collect and use data. However, she is willing to pay the price given the increased autonomy she has in running her school. An issue is that sometimes she feels that the system is 'awash with data'. Over time she has learnt to make the best of information that is available to 'make sense' of initiatives that have been implemented in the school. From time to time it has been necessary to plan small internal evaluation studies for which data collection methods had to be developed, rather than relying on existing data sources. She sees this as a way of developing and extending the notion of a learning culture across the school.

5. Comparisons with Theory

Key strategies used in the case example have been outlined in this paper. I now reflect on them in terms of current organisational and evaluative principles.

1) Control of the WSE Framework. Evaluation objectives and priorities are set by the system, not the School. Conventional evaluation practice begins with the identification of issues that need to be investigated (or

evaluation questions). These are often the objectives of the intervention being studied. Evaluators then look for appropriate data sources in order to answer these questions. The situation is reversed in WSE in Victoria. Schools are provided with a range of data in the form of indicators and must 'work backwards' for issues that relate to them. Implicit in the data provided are the priorities of the school system, for example school performance in numeracy and literacy.

2) Evaluation as Performance Management. Conventional impact evaluation looks for causal relationships between processes and outcomes. The evaluator looks for variables that link cause and effect. WSE in Victoria separates cause and effect. The central authority provides a series of indicators that provide systematic information on school performance, largely student outcomes. Decisions about process are the responsibility of the school, for example how to organize the teaching of literacy. In the Review process it is the professional knowledge of school leaders, teachers and external reviewers that provide the data for these decisions. Cause and effect are effectively separated and rely on different forms of data. This approach is consistent with a performance management approach to evaluation.
3) WSE as Perceived by Principals. As implied in the case study above, WSE has meant an additional load for school leaders. However, Principals that have now participated in several rounds of the School Review process believe that the extra work load has been worth the effort. Of particular importance is that school leadership knows where they stand in terms of relative school performance. This is important information in setting school directions, and in internal decision-making. Whereas, in the past opinions of staff often swayed such decisions, a Principal can now point to systematic findings about core objectives in debates about school priorities. It is good to see that in the most recent version of the Strategic Planning documentation that the DoE is providing support for the implementation of WSE via the creation of Senior Education Improvement Leaders.
4) Evaluation in Non-Priority Curriculum Areas. While school evaluation is closely linked to system priorities, it is possible for enlightened principals to commission studies that respond to a local rather than system requirements. It may well be proactive, for example to assess the extent of an educational need. This might relate to an area of curriculum that is not a system level priority, or might reflect an external influence due to a pressure group. For example, the need to develop a program designed to improve student health, through walk-to-school or bike-ed strategies. However, there is no indication from the DoE indicating encouragement or support for such endeavours.
5) WSE and Accountability: The Recent Past. Accountability has been defined as the responsibility for the justification of expenses, decisions or results of one's own efforts. It is often said that school leaders and teachers should be accountable for their pupils' achievements. Accountability can be thought of as a transaction along the lines of 'if we give your resources, we expect you to show what you have done with them'. An implication is that sanctions may be applied. This could come in the form of reprimand, a note on your file, or even the loss of a job. In Victoria, there has been a long tradition of devolving responsibility to schools for the way they deliver educational services to their local community and a one-line budget. However, this is within guidelines set by a central body, now the DoE, including curriculum priorities such as literacy and numeracy. It is fair to say that the original objective of WSE was to serve the needs of schools and school principals to feed into what I will call an *Accountability UP* process. In the conduct of *Accountability UP* schemes Performance Management has been championed as an efficient way of assembling and disseminating simple but understandable information for reporting to the next level up in a systemic hierarchy. The DoE in Victoria requires such information to make statements to government and the public about the general health of the government system. In addition, large scale outcome data has

enabled policy makers to identify lower performing schools, and to apply remediation where possible. This might also involve actions such as the removal of a school principal and/or result in closure or amalgamation of schools.

6. Conclusion

The current WSE scheme in Victoria still acknowledges the need to feed into an *Accountability UP* perspective. However, as the Review framework has matured, there has been an increased emphasis for DoE support for school improvement via successive rounds of Strategic Plan development. School Reviews now encourage the use of WSE for enhancing school improvement and performance.

It can be argued that WSE in Victoria now meets the information needs of both the DoE *and* individual schools. One can think of this in terms of a two-level system of mutual support. WSE feeds the information needs of both the DoE and each individual school.

One can think of these arrangements in terms of *Accountability UP and Accountability DOWN* (Owen 2007).

This has been brought about by a combination of factors, including; the sympathy of most external reviewers towards the information needs of the schools, and the provision of indicator information that is meaningful to principals and school councils. This includes information that allowed them to compare the performance of their school with 'like schools', those with similar student populations on socio-economic grounds.

In pragmatic terms, the current framework could be regarded as a satisfactory arrangement in which evaluation plays an influential role.

References

Earl, L. and S. Katz (2010), Creating a Culture of Inquiry: Harnessing Data for Professional Learning. In Blankstein, A., Houston, P. and Cole, R. (eds.) *Data Enhanced Leadership (The Soul of Educational Leadership Series)*, Corwin Press, Thousand Oaks, CA, 94-116.

Ikin, K., & McClenaghan, P. (2015). Modelling the Influences of Evaluation on School Principals. *Evaluation Journal of Australasia*, 15 (1), 19-27.

Owen, J.M. (2007). Making Policy Interventions More Effective: The Case for Accountability Up and Accountability Down. In Bemelmans-Videc, M. et.al. *Making Accountability Work: Dilemmas for Evaluation and for Audit*. Transaction Publishers, New Brunswick NJ, 181-192.

Sanders, J.R., & Davidson E.J. (2003). A Model for School Evaluation. *International Handbook of Educational Evaluation*, 807-826.

(Accepted 17th September, 2017)

【Research Note】

School Evaluation in Korea: Continuities, Changes and Challenges[1]

Sung Jae PARK
Korean Educational Development Institute
sjpark@kedi.re.kr

Abstract

This paper reviews the school evaluation system in Korea to identify the system's characteristics and problems for provision of a new direction. The jurisdiction on schools is prescribed by law and the school evaluation is a part of administrative affairs. Accordingly, school evaluations over elementary and secondary schools are controlled by the 17 local autonomous entities and higher educational institutions overseen by the central government (Ministry of Education). In the case of elementary and secondary schools, the evaluations are further diversified depending on the elected local superintendent. Recently, external evaluation has mostly disappeared as it was replaced by internal evaluation and follow-up consultation that was suggested as alternatives to improve the educational quality. Although quantitative evaluation as prescribed by law is mainly used in the evaluation process, qualitative evaluation is what actually determines the results, and typical perspective has been transformed from "positivistic" to "constructivist" paradigm. In this process, there has been an emerging issue on the accountability of school education. To review those matters, this paper considered first, the relationship between school evaluation system and the local governments, and summarized periodical transition process of evaluation system. And finally a comparative analysis was conducted to study the managing status of school evaluation in each local government body by studying the 16 MPOEs except one, the newly opened Sejong Office of Education.

Keywords

School Evaluation, Internal Evaluation, External Evaluation, Positivistic Paradigm, Constructivist Paradigm

Introduction

The evaluation process of elementary & secondary schools and higher educational institutions are basically similar. Elementary and secondary schools are evaluated every year and the Metropolitan and Provincial Offices of Education (MPOEs) review and decide the need for school consultation based on the evaluation results or by request of schools that are willing to change. Then, the MPOEs commission the consulting project to Korean Educational Development Institute (KEDI), and KEDI forms expert groups to carry out school consultations. In most cases, after

the consulting is done, the results are then shared with the school and the MPOE for financial support and administrative measures. However, this was the previous evaluation scheme conducted in 2010-2014. As school evaluation is delegated to the MPOEs, most of the quantity-based evaluation and peer reviewed external evaluation were replaced by internal evaluation and follow-up consulting procedures, which would simplify the whole evaluation process and focus more on improving the school in question. The following are the details of the evaluation, such as evaluation procedures, indicators, contents, methodology and problems.

School Evaluation in Korea

1.1. Overview

There are three types of school evaluation in Korea; external evaluation, external & internal evaluation and internal evaluation. The MPOEs provide consulting services to the schools with poor evaluation result. Though the Ministry of Education (MOE) revised the law to give the MPOEs the autonomy in school evaluations, it should be noted that the revision only gave partial discretion in operating the measures regulated by the areas and procedures in the statute. The Enforcement Decree of the Elementary and Secondary Education Act Article 12 (Criteria of School Evaluation) Section 2 defines three implementation areas of school evaluation as follows: (i) operating curriculum and teaching & learning; (ii) educational activity and achievement; and (iii) other areas on school management authorized by the chief of the MOE or a superintendent. In other words, implementation areas are divided into essential and additional sections, and the superintendent only has discretion on the latter. Article 13 (procedure and announcement of evaluation) Section 3 states that school evaluations are basically of quantity-based, but if quantitative evaluation is insufficient, qualitative evaluation such as documentary evaluation, survey and counseling are allowed. Although the law defines qualitative evaluation as optional, school evaluation is implemented differently by each MPOE. As educational administration becomes more decentralized, school evaluations sometimes depend on political agenda of the elected superintendent and deviate from the intent of the law. In the following sections, this paper compares the school evaluations of different timelines and operative methods of the 16 MPOEs in a specific year.

<Periodic Changes in School Evaluation>

1. Preparation Phase for School Evaluation: 1996-1999

In 1995, Presidential Committee on Educational Reform (PCER) marked the importance of the school evaluation system in [Educational reform measures to establish new education system] and made evaluation results consequential to government's financial support. In 1996, the Ministry of Education (MOE) added school evaluation to the performance appraisal standards of the Metropolitan and Provincial Offices of Education (MPOEs), triggering all education offices to embark on nationwide school evaluation. In 1997, Elementary and Secondary Education Act Article 9(2) stated legal grounds for national school evaluation. In 1998, the Enforcement Decree of the Elementary and Secondary Education Act Article 11 and 12 stipulated subjects and standards of the school evaluation system. The MOE's [The Five-year Plan for Educational Development] of 1999 suggested implementation of measures for school evaluation. The early stage of school evaluation was a preparation period, significant in laying the legal groundwork for the school evaluation system.

2. Introduction of National-level School Evaluation: 2000-2003

From 1998 to 1999, KEDI developed the school evaluation system and verified the validity of the evaluation

model through test operation. After undergoing a series of test evaluations between 2000 and 2001, the actual school evaluation took place during 2002 to 2003. The school evaluation system on a national level began to materialize through this process. This is the period when, besides the national-level evaluation, the MPOEs came up with their own plan to conduct school evaluations and carried them out with regular intermissions of 1 to 3 years. During this period, dual enforcement of school evaluations were enacted – one on a national level and the other on a metropolitan and provincial level.

3. Implementation Period of School Evaluation using Common Indicators: 2004-2010

Under the dual enforcement of school evaluation system, the central government (MOE) provided the MPOEs with common indicators to conduct school evaluations. First of all, on the national level, the central government developed and supplied common evaluation indicators through design, training and monitoring. At the metropolitan and provincial levels, the MPOEs carried out the actual evaluation while using both the common indicators from the central government and their own indicators. The school evaluation of this period was funded by shared expenses of the MPOEs, so it was possible for them to pursue with more autonomy.

4. Autonomous Evaluation by the MPOEs: 2011-Present

The MPOEs are delegated legal basis for autonomy as the Enforcement Decree of the Elementary and Secondary Education Act was revised on February, 2013. The common indicators developed on the national level were provided as a guideline, as the choice of whether or not to use the indicators was based on the autonomous decision of the MPOEs. The Enforcement Decree of the Act states that only the indicators in evaluation and quantitative evaluation principle (qualitative evaluation is exceptionally admitted here) so that each of the 17 MPOEs could obtain their autonomy. Expanded autonomy of local education offices (MPOEs) marked a milestone in this period, because each office conducted evaluation under the regional specifications and educational conditions. Particularly since 2014, superintendents of the MPOEs offered their own policies, which completely shifted the system from external evaluation to each school's self-evaluation. The most significant change in this period is the rapid alteration of quantity- and relative-oriented evaluation into quality & non-measure evaluation for autonomous delegation by the 17 MPOEs. According to many progressive superintendents (13 of 17) who are elected in 2014 local election, they have been trying to move away from authoritarian and regulative evaluation. Politically, these elected superintendents have reasons to respond to the teachers' negative perception on school evaluation and voices calling for the abolition of numerous evaluations. This is also the time when the MOE newly instituted university evaluation. The MOE planned to restructure universities from year 2014 to 2023. The first phase of university evaluation (from 2014-2017) is now in progress. It aims to reduce the number of universities (currently there are over 400) and improve the quality of education to prepare for a decrease in population, changes in the labor market, and emergence of artificial intelligence.

Source: KEDI School Evaluation Website: http://eval.kedi.re.kr/rspage.jsp?mn=2&sm=201; School Evaluation Guidebook 2010-2016, KEDI; [Supplemented by the author]

1.2. Methodology of the Analysis

1.2.1. Reference Point and Data

Reference point of the analysis is 2012 and 2016, when the superintendents of 2010 and 2014 elected through the local elections met their second year in office; a policy would mostly be effective after two years of its adoption. The Office of Education in Sejong city (metropolitan city) is excluded as the office is newly instituted. This paper

reorganized the data which are the Basic Plan for annual School Evaluation submitted by the 16 MPOEs.

1.2.2. Analytic Framework

The analytic framework of this paper refers to Y. K. Lim (2005: 162-163)'s school evaluation models and features. This paper then compares the school evaluation policies of 2010 superintendents to that of 2014 superintendents. The following are basis derived from major issues and perspectives on school evaluation to analyze each model and feature.

School evaluation has long been under discussion in the evaluation literatures (Scriven, 1967, 1991; Stufflebeam et al., 1971, 2014; Love, 1991; Sonnichsen, 2000; Nevo, 1995, 2002). The meaning of school evaluation differs by perspectives. According to David Nevo (1995, 1998), Tyler (1950) saw school evaluation as a process to identify the degree of realizing of an educational goal; Stufflebeam (1969) and Alkin (1969) saw it as offering of information for decision making; and Stufflebeam (1974), Eisner (1979) and House (1980) saw it as an inspection for valuation. Guba and Lincoln (1989: 21-49) categorized views on evaluation into three generations and continued on to suggest the fourth generation evaluation. The first three generations are 'measurement', 'description' and 'judgement' by order, and the authors claim that they are based on scientific paradigm of an objective substance. Guba and Lincoln's fourth generation evaluation negates scientific paradigm and defines it as a discursive and negotiable process on the basis of pluralism and constructivism (Laughlin and Broadbent, 1996). Here, evaluation is to identify issues on parties of interest, review the problems through negotiation between evaluator and other parties of interest and seek ways to improve school education. The concerned parties are subjects who construct their own world of defining matters. Intervention by an evaluator is unnecessary if such construction is accomplished by negotiation. In short, there are two kinds of perspectives on evaluation. One is 'Offering information for decision making' and 'Inspection for valuation' based on scientific paradigm, and the other is 'evaluation as discussion and negotiation' incorporating the offering of information and judging values on the basis of responsive constructivist paradigm (Guba and Lincoln, 1989: 43-49; Lim, 2005: 162-163). Issues on school evaluation are separated by 5W1H: Who evaluates on school? (Subject); When are evaluations executed? (Period and Schedule); Where does the evaluation take place? (Place); What is the evaluation's objective? (Objective); Why does the evaluation is executed? (Purpose); and How can we evaluate schools and applicate result? (Methodology and Application) Keeping these standards in mind, the followings are comparisons on management status on school evaluation in the 16 MPOEs.

Table 1 Analytic framework

	Category	Standard
Perspectives	Paradigm	Scientific vs. Constructivism
Issues	Purpose	Accountability vs. Improvement
	Subject	External Evaluation vs. Internal Evaluation
	Period	Every Two to Three Years vs. Every Year
	Methodology	Quantitative vs. Qualitative
	Application	Compensation vs. Searching for Alternatives

Source: Y. K. Lim (2005: 169) [reconstructed by the author]

1.3. Analysis on School Evaluation in the 16 MPOEs

1.3.1. Purpose

Reviews on school evaluation of each MPOEs under the administration of elected superintendents in 2010 show that the purpose emphasizes on accountability. In addition to this, these MPOEs suggest supplementary purposes along with the regional characteristics. Most of them put emphasis on accountability, improvement of educational quality and policy, and self-diagnosis. The 2010 period's purpose places more weight on educational accountability than autonomy of school management (see Table 2 (left)).

Administration in 2014, purposes on evaluation that stresses autonomy (13), improvement of educational quality (13), accountability (10), focusing on regional education (7), policy making and development (6) (see Table 2 (right)). In other words, they intend to improve educational quality through autonomous school management. The issue is how to secure accountability on the situation relying on internal evaluation.

1.3.2. Subject and Methodology

The Enforcement Decree of the Elementary and Secondary Education Act Article 13(2) of 2010 states that school evaluation may have document evaluation, field evaluation and comprehensive evaluation which incorporate both internal and external aspects. Additionally, it should inquire responses of students and parents through various means such as surveys and interviews to include them in evaluation. Therefore, the year 2010 evaluation integrates internal

Table 2 Purpose of Evaluation in the 16 MPOEs

	Purpose of Evaluation in 2011							Purpose of Evaluation in 2016						
	Autonomy	Improvement	Accountability	Customer Satisfaction	Sharing Best Practice	Policy Making	Self-Diagnosis	Autonomy	Improvement	Accountability	Customer Satisfaction	Sharing Best Practice	Policy Making	Focusing on Regional Education
Seoul	O	O	O			O		O	O				O	
Busan	O	O	O		O	O	O	O	O	O	O	O	O	
Daegu		O	O		O	O	O	O	O					O
Incheon	O	O	O			O		O	O					O
Gwangju	-	-	-	-	-	-	-		O			O	O	O
Daejeon	-	-	-	-	-	-	-	O	O	O		O		
Ulsan	O	O	O		O		O	O	O	O			O	
Gangwon	O	O	O		O	O		O	O	O			O	
Gyeonggi		O	O			O	O	O	O	O				
Chungbuk	-	-	-	-	-	-	-	O	O	O				
Chungnam	-	-	-	-	-	-	-		O	O	O			
Jeonbuk		O	O				O	O	O	O				O
Jeonnam			O			O	O	O	O	O			O	
Gyeongbuk			O	O			O	O			O			O
Gyeongnam	O	O			O	O								O
Jeju			O				O	O		O				O
Total	6	9	11	1	5	8	8	13	13	10	3	3	6	7

Source: KEDI [reconstructed by the author]; The 16 Metropolitan and Provincial Offices of Education

※Based on Basic Plan for School Evaluation, some MPOEs are excluded due to unclear purposes.

※Due to the material submitted, 'Purpose of Evaluation in 2011' is used instead of that of 2012.

and external evaluations, and it takes the form of internal evaluation, document evaluation and visiting evaluation. In internal evaluation, every school should hold group discussion and submit evaluation report. Then the MPOE forms an external evaluation team, and they implement visiting evaluation after document evaluation and submit an evaluation report. Document evaluation and visiting evaluation, to alleviate burden of evaluation, are decreasing in number since 2011. In 2012, MPOEs of Gangwon and Gyeonggi implemented internal evaluation without an external one, and Gyeongbuk MPOE conducted only quantitative evaluation (see Table 3 (left)).

On the other hand, the most significant change in 2014 administration of elected superintendent was on the subject of, "who evaluates the schools." Article 13 has been revised, but that was not the only reason for the newly

Table 3 Evaluation Methodology

	Evaluation Methodology in 2012				Evaluation Methodology in 2016	
	Internal Evaluation	Visiting Evaluation	Document Evaluation	Remarks	Internal Evaluation	External Evaluation
Seoul	O	O	O		O	
Busan	O	O	O		O	
Daegu	O	O	O	Visiting Evaluation 1/3 in each level of school		O
Incheon	O	O	O		O	
Gwangju	O				O	
Daejeon	O	O	O	Visiting Evaluation on upper 15% Schools		O (Connected to School Management and MPOEs Evaluation)
Ulsan	O		O	Satisfaction Survey on entire School	O	O
Gangwon	O			Internal Evaluation Only	O	
Gyeonggi	O			Internal Evaluation Only	O	
Chungbuk	O	O	O		O	
Chungnam	O	O	O		O	
Jeonbuk	O	O			O	
Jeonnam	O	O	O	Minimizing Visiting Schools	O	
Gyeongbuk				Quantified Evaluation Only.	O	O
Gyeongnam	O	O			O	
Jeju	O	O	O		O	
Total	15	11	10		14	4

Source: KEDI [reconstructed by the author]; The 16 Metropolitan and Provincial Offices of Education

elected superintendents avoiding external evaluation and changing to an internal one. Evaluators also changed from an outsider to insider. In 2016, most of MPOEs are only implementing internal evaluation, except for few cases like Ulsan and Gyeongbuk where MPOEs conduct both internal and external evaluations and Daegu and Daejeon carrying out only external evaluation (see Table 3 (right)). Without external evaluation, however, it is difficult to guarantee objectivity and impartiality. Since internal evaluation is a form of self-evaluation by school members, it may omit social agreement on detailed standard, indicators and reasonability.

1.3.3. Indicators of School Evaluation

Yearly indicators on school evaluation in 2010 administration are as follows. In 2010, indicators are divided into common indicators of the national level and individual indicators of MPOEs. Scores on individual indicators is autonomously regulated by the MPOEs. Common indicators, applied nationally without any modification in every MPOE consist of 14 contents; educational goal (2), education process and method (4), educational achievement (4) and educational management (4). In 2011, evaluation followed standard suggested by the Enforcement Decree of the Elementary and Secondary Education Act, therefore, it included four areas; educational process & teaching/learning, educational management, educational achievement and satisfaction. The emphasis is on achievement-oriented evaluation indicators rather than existing activity-oriented evaluation system. There are 22 indicators that consist of education process & teaching/learning (2 quality based indicators of 7), education management (1 quality based indicator of 5), education achievement (9), and satisfaction (1). The MPOEs can decide the implementation area and weight of each indicator. In 2012, indicators were assigned by the MPOEs autonomously. The 2012 school evaluation indicators provided by KEDI was a guideline on evaluation indicators to the MPOEs rather than common indicators of the national level. As the autonomy of the MPOEs gained strength, national evaluation indicators in 2010 transformed into standardized indicators provided by the central government (refer to the KEDI School Evaluation Guidebook 2010-2012). In 2012, quality indicators and quantity indicators were applied at the same time but Gwangju, Ulsan and Gyeongbuk used only quantity indicators (see Table 5 (left)). All the MPOEs use both the

Table 4 Yearly comparison to school evaluation indicators in 2010-2012

	2010	2011	2012
Type	Common indicators in national level & Internal indicators in MPOEs	Common indicators in national level & Internal indicators in MPOEs	Common indicators in national level & Internal indicators in MPOEs
Methodology	Quality indicators	Quality indicators and Partial quantity indicators	Quantity indicators
Area	1. Education goal 2. Education process and method 3. Education achievement 4. Education management	1. Education process & teaching/learning 2. Education management 3. Education achievement 4. Satisfaction	1. Education process & teaching/learning 2. Education management 3. Education achievement 4. Satisfaction
Application of indicators in national level	Common indicators applied nationally without any modification in every MPOE	Autonomously assignment to the proportion of area and weight of indicator by MPOEs.	Autonomously assignment to selection, application, proportion of area and weight of indicator by MPOEs.

Source: KEDI [reconstructed by the author]; The 16 Metropolitan and Provincial Offices of Education

quantity-oriented evaluation made by the MPOEs and the indicators provided by KEDI, except for Gyeongbuk MPOE which only applies the KEDI indicators. The portion of applying internal indicators in the MPOEs is gradually increasing. It is positive that indicators developed by the MPOEs considering regional characteristics are more suitable in field evaluation. However, there are political concerns that school evaluation is used for favorable valuation on key projects of superintendents of the MPOEs.

In 2014 administration, most of the MPOE applies quantity and qualitative evaluation although the quality evaluation is only partially used. However, the proportion of scoring in qualitative evaluation or internal evaluation developed for the major policy of superintendent is also growing. Daegu MPOE applied absolute evaluation (100% quantitative evaluation). Jeju used quantitative evaluation for common indicators and chose qualitative evaluation as optional (internal) indicators. In addition, most of the MPOEs employed both common and internal indicators at the same time. Daejeon and Ulsan only used common indicators. Daegu, Chungbuk and Gyeongbuk applied common indicators, optional indicators and autonomous indicators altogether (see Table 5 (right)).

In 2010 administration, school evaluations were executed as prescribed by the Enforcement Decree of the Elementary and Secondary Education Act Article 13(2). However, selecting indicators, application, proportion of indicator area and weight on each indicator were autonomously determined by the MPOEs (see Table 6 (left)).

Table 5 Forms and configuration of school evaluation indicators in the MPOEs

	Forms and configuration of school evaluation indicators in 2012				Forms and configuration of school evaluation indicators in 2016					
	Forms of indicators (Quantity indicator base)		Configuration of indicators		Forms of indicators		Configuration of indicators			
	Quantity Indicators	Quantity + Quality indicators	Common Indicators (KEDI)	Internal Indicators (MPOE)	Quantity Indicators	Quantity + Quality Indicators	Common Indicators (KEDI)	Internal Indicators (MPOE)	Optional Indicators	Autonomous Indicators
Seoul		O	O	O		O	O			O
Busan		O	O	O		O	O			O
Daegu		O	O	O	O		O		O	O
Incheon		O	O	O		O	O			O
Gwangju	O		O	O		O	O	O		
Daejeon	O	O	O	O		O	O			
Ulsan	O		O	O	O		O			
Gangwon		O	O	O		O	O	O		
Gyeonggi		O	O	O		O	O			O
Chungbuk		O	O	O		O	O		O	O
Chungnam		O	O	O		O		O		O
Jeonbuk		O	O	O		O	O	O		
Jeonnam		O	O	O		O	O		O	
Gyeongbuk	O		O		O		O		O	O
Gyeongnam		O	O	O		O	O			O
Jeju		O	O	O		O	O			O
Total	4	13	16	15	3	13	15	4	4	10

Source: KEDI [reconstructed by the author]; The 16 Metropolitan and Provincial Offices of Education

Table 6 Area and management of school evaluation indicators in the MPOEs

	Area and management of school evaluation indicators in 2012		Area and management of school evaluation indicators in 2016	
	Area of school evaluation indicator (legal statement)	Management of School evaluation indicator	Area of school evaluation indicator (legal statement)	Management of School evaluation indicator
The Enforcement Decree of the Elementary and Secondary Education Act Article 13(2)	② School evaluation is implemented as stated by following statement 1. Managing educational process and teaching/learning method 2. Educational activity and achievement 3. Matters for school management admitted by chief of the MOE or superintendent		② School evaluation is implemented as stated by following statement 1. Managing educational process and teaching/learning method 2. Educational activity and achievement 3. Matters for school management admitted by chief of the MOE or superintendent	
KEDI Guideline Indicator	1. Educational process and teaching/learning method, 2. Education management, 3. Educational achievement, 4. Satisfaction	Autonomously assign to selection, application, proportion of area and weight of indicator by the MPOEs.	1. Educational process and teaching/learning method 2. Education management 3. Educational achievement 4. Satisfaction	Autonomously assign to selection, application, proportion of area and weight of indicator by the MPOEs.
Seoul	Same as above	Same as above	1. Educational vision of Seoul, 2. Educational process and teaching/learning, 3. Educational activity and achievement	Same as above
Busan	Same as above	Same as above	1. Educational process and teaching/learning, 2. Education management, 3. Educational achievement	Same as above
Daegu	Same as above	Same as above	1. Educational process and teaching/learning, 2. Education management, 3. Educational achievement, 4. Satisfaction	Same as above
Incheon	Same as above	Same as above	1. Secure and peaceful school, 2. Education on creativity and sympathy, 3. Education welfare for all, 4. Fair and transparent educational administration	Same as above
Gwangju	Same as above	Same as above	1. Educational process and teaching/learning, 2. Education management, 3. Educational achievement	Same as above
Daejeon	Same as above	Same as above	1. Educational process and teaching/learning, 2. Education management, 3. Educational achievement, 4. Extra points, 5. School evaluation report (Qualitative evaluation)	Same as above
Ulsan	Same as above	Same as above	1. Educational process and teaching/learning, 2. Education management, 3. Educational achievement, 4. Satisfaction	Same as above
Gangwon	Same as above	Same as above	1. Educational process and teaching/learning, 2. Education management, 3. Educational achievement, 4. Satisfaction	Same as above
Gyeonggi	Same as above	Same as above	1. Self-governance community of participation and communication, 2. Lifestyle community of respect and care, 3. Learning community of openness and cooperation, 4. Managing creative education process	Same as above
Chungbuk	Same as above	Same as above	1. Making democratic education culture, 2. Constructing school system focusing on educational activity, 3. Managing creative education process, 4. Satisfaction on education process	Same as above
Chungnam	Same as above	Same as above	1. Educational process and teaching/learning, 2. Education management, 3. Educational achievement, 4. Satisfaction	Same as above
Jeonbuk	Same as above	Same as above	1. Educational process and teaching/learning, 2. Education management, 3. Educational achievement, 4. Education policy of Jeonbuk	Same as above
Jeonnam	Same as above	Same as above	1. Educational process and teaching/learning, 2. Education management, 3. Educational achievement, 4. Satisfaction	Same as above
Gyeongbuk	Same as above	Same as above	1. Educational process and teaching/learning, 2. Education management, 3. Educational achievement, 4. Satisfaction	Same as above
Gyeongnam	Same as above	Same as above	1. Education philosophy, 2. Educational process and teaching/learning, 3. Education activity, 4. Autonomous management system	Same as above
Jeju	Same as above	Same as above	1. Educational goal, 2. Educational process and method, 3. Educational achievement, 4. Education management	Same as above

Source: KEDI [reconstructed by the author]; The 16 Metropolitan and Provincial Offices of Education

In 2014 administration, most of the MPOEs set evaluation area as the Enforcement Decree states, and some MPOEs include evaluation area suitable to their regional characteristics (see Table 6 (right)). Most of evaluation indicators consist of common and autonomous indicators, quantitative evaluation and qualitative evaluation all at the same time. Exemplary materials were provided for application of indicators to schools in case of autonomous indicators. But evaluation areas of some MPOEs totally derailed from standard stated in the law, and even if standard was fixed in accordance with the law, there were cases that scored on specific policies promoted by superintendent held more weight in the policy portfolio. Therefore, some experts point out that school evaluation is misused considering the original intent due to political motivation of superintendents.

1.3.4. Implementation Period

In 2010 administration, school evaluation was implemented to all national, public, and private level of elementary, secondary, and special schools stated in the Enforcement Decree of the Elementary and Secondary Education Act Article 11 (on the subject of evaluation) based on the MOE's Basic Plan for School Evaluation. School evaluation was already initiated in 2006 and took place every three years from then. The 2006-2008 period was the first phase, and the 2009-2011 phase was the second. As the MPOEs were free to select the subjects of evaluation, many MPOEs chose 1/3 of all elementary, secondary, and special schools under their jurisdiction. In 2012, 9/16 of the MPOEs implemented school evaluation every year (see Table 7 (left)).

In 2014 administration, most MPOEs executed school evaluation every year. Daejeon MPOE implemented

Table 7 School evaluation implementation period in the MPOEs

	School evaluation implementation period in 2012				School evaluation implementation period in 2016		
	Every year	A grade of school In each year	Every two year	1/3 schools In each grade of school In each year	Every year	Every two year	1/3 schools In each grade of school In each year
Seoul				O	O		
Busan				O	O		
Daegu	O				O		
Incheon	O				O		
Gwangju	O				O		
Daejeon	O (Elementary & High Schools)			O (Middle Schools)	O (Elementary & Middle Schools)		O (High Schools)
Ulsan	O				O		
Gangwon	O				O		
Gyeonggi	O				O		
Chungbuk		O			O		
Chungnam	O				O		
Jeonbuk				O	O		
Jeonnam				O	O		
Gyeongbuk	O				O		
Gyeongnam			O		O		
Jeju				O		O	
Total	9	1	1	6	15	1	1

Source: KEDI [reconstructed by the author]; The 16 Metropolitan and Provincial Offices of Education

frequent evaluation (quantity / every year) to elementary and secondary (middle) schools when comprehensive evaluation was absent and conducted comprehensive evaluation (quantity and quality / every three year) to secondary (high) and special schools. Jeju MPOE executed evaluation every two years, considering the size of school and regional characteristics. Generally, an evaluation is scheduled from March to February of the next year, but it slightly differs by evaluation indicators and the MPOEs (see Table 7 (right)).

1.3.5. Employing School Evaluation Result

In 2010 administration, school evaluation results were utilized as the following (see Table 8 (left)). The results are uploaded on their school homepage focusing on outstanding performance and suggestion. The MPOEs recommended school consultation with tailored administrative and financial support based on the results. The MOE, which led school evaluation, shared best practices with entire schools to contribute to the development of school education. Main characteristic of this period was the building of plans to execute follow-up consultation in every MPOE from 2012.

In 2014 administration, the MPOEs uploaded school evaluation result on their school homepage and website to notify the public on school information. Only those external evaluations used by the MPOE, in Daegu, Daejeon, Ulsan and Gyeongbuk, classified the evaluation results, and provided administrative and financial supports upon the results (see Table 8 (right)). This was a big difference from the 2010 superintendent administration. On the feedback process, analysis on result was applied to follow-up consultation and employed as materials **to build a school education plan** in most of the MPOEs.

Table 8 Employing school evaluation result in the MPOEs

	Employing school evaluation result in 2012					Employing school evaluation result in 2016				
	Follow-up Consultation	Administrative & Financial Support	Giving an award for best school	Giving an award for best teacher and HR benefit		Follow-up Consultation	Administrative & Financial Support	Giving an award for best school	Giving an award for best teacher and HR benefit	Release to website for public notice on school information
Seoul	O	O	O			O				O
Busan	O	O	O	O		O				O
Daegu	O	O	O	O		O	O			O
Incheon	O	O	O			O				O
Gwangju		O	O			O				O
Daejeon	O	O	O	O		O	O	O	O	O
Ulsan	O	O	O	O		O	O	O	O	O
Gangwon	O	O				O				O
Gyeonggi	O	O				O				O
Chungbuk	O	O	O			O				O
Chungnam	O	O	O	O		O				O
Jeonbuk	O					O				O
Jeonnam	O	O	O	O		O				O
Gyeongbuk	O	O	O			O	O			O
Gyeongnam	O		O	O		O				O
Jeju	O	O	O	O		O				O
Total	15	14	13	8		16	4	2	2	16

Source: KEDI [reconstructed by the author]; The 16 Metropolitan and Provincial Offices of Education

Discussion and Conclusion

School evaluation in Korea is changing as external evaluations disappeared and were replaced with internal evaluations. Along the way, weighing the values of accountability and autonomy became a controversial issue. It implies that school evaluation is altering from the external evaluation controlled by the outside institutions to internal communicative model. Application of external evaluations lessened as school consultation and diagnosis on school organization have expanded. While school evaluations that are usually harsh on teachers disappeared, school consultation and organizational diagnosis strengthened, because they engage in participatory process of discussion and negotiation to solve their problems.

Some of the issues on school evaluation include; (i) The subject of evaluation - it has altered from external evaluation to internal evaluation and main value also changed from effectiveness-centered evaluation to validity-centered evaluation, but the problem of accountability still remains. (ii) The evaluation tools - main indicators are qualitative, not quantitative. Although the law specifies main indicators as quantitative evaluation, and qualitative evaluation as supplementary, the proportion in application is to be autonomously decided by the MPOEs, so now qualitative evaluation holds more weight than quantitative evaluation. (iii) The perspective on evaluation has changed from positivistic paradigm to constructivist paradigm, which focuses on self-problem-solving consultation and organizational diagnosis in cooperation with outside experts. (iv) The correlation of school evaluation result with another performance evaluation - it was found that school evaluation result was related to principal evaluation, performance-based pay and individual HR assessment in the past, but not anymore. (v) The focus on evaluation has changed from result-driven evaluation to follow-up consultation which aims for educational improvement. (vi) The purpose of evaluation adjusted from accountability to autonomy or solving the problems. (vii) The effectiveness of evaluation valued on partnership and cooperation rather than competitiveness. (viii) The changes of school evaluation also influenced the elementary and secondary school from running an examination-oriented curriculum to a characterized program. (ix) In regards to the purpose of the evaluation, it also shifted from assessment-, description-, and judgment-oriented one to a more communicative and collaborative one where the traces of fourth generation evaluation of Guba & Lincoln (1989) can be found.

But problems remain: (i) Confusion in transition period or early stage of introduction are expected. It needs plenty of time to examine and adjust to a new system. (ii) Despite decent intention of autonomous evaluation by the MPOEs, present system has a possibility of causing the dissolution of evaluation system that is prescribed by the law. Although law regulates the areas of school evaluation indicators and sticks to quantitative evaluation principle, qualitative indicators determine evaluation result by adjusting proportion between quantitative and qualitative indicators in the MPOEs. In addition, with only one internal evaluation process, it is insufficient to evaluate a complex entity such as school, hospital and army, so the school evaluation could be deviating from the intent of the law. (iii) Accountability is difficult to measure in the context of "Value for Money." The current school evaluation system is impossible to compare or crosscheck the results among the MPOEs or schools. For school evaluations are internal evaluation, it is hard to comprehensively understand the entire school system from an objective point of view. The current school evaluation system may be problematic in that educational and financial accountability cannot be secured among schools that are running on taxpayers' contributions. And finally, (iv) there could be further questions on the effectiveness and efficiency of the current evaluation system.

Note

1 This article is a revised version of author's conference presentation at the Japanese Evaluation Society held on November 26, 2016.

References

Alkin, M. C. (1969). Evaluation theory development. *Evaluation Comment*, 2, 2-7.

Buckley, K. & Marion, S. (2011). *Approaches and considerations for Incorporating Student Performance Results from "Non-Tested" Grades and Subjects into Educator Effectiveness Determinations,* 3-5. www.nciea.org

Eisner, E. W. (1979). *The educational imagination*. New York: Macmillan.

Gu, J. O., Park, S. J., Ahn, B. H, Lee, S. R. and Jung, T. H. (2013), *School Consulting: Theory and Practice*. Korean Educational Development Institute. Seoul.

Guba, E. G. & Lincoln, Y. S. (1989). *Fourth generation evaluation*. Newbury Park, CA: Sage.

House, E. R. (1980). *Evaluating with validity*. Beverly Hills, CA: Sage.

KEDI (2017) "School Evaluation System": http://eval.kedi.re.kr/rspage.jsp?mn=2&sm=201

Laughlin, R. & Broadbent, J. (1996). Redesigning Fourth Generation Evaluation: An Evaluation Model for Public-Sector Reforms in the UK? *Evaluation* 2(4): 431–51.

Lim, Y. K. (2005). The Pattern and Characteristics of School Evaluation in Korea. *The Journal of Educational Administration* 23(1).

Love, A. J. (1991). *Internal evaluation: Building organizations from within*. Newbury Park, CA: Sage.

Nevo, D. (1995). *School-Based Evaluation: A Dialogue for School Improvement*. Oxford: Pergamon.

Nevo, D. (1998). Dialogue evaluation: a possible contribution of evaluation to school improvement. In J. C. Tedesco (Ed.), Prospects. *Quarterly review of comparative education. The evaluation of education systems: the view at the end of the 1990s*. Vol. XXVIII, no. 1. Paris: UNESCO. March 1998.

Nevo, D. (2002). *School-based Evaluation: An International Perspective.* JAI: Elsevier Science.

OECD (1995). Schools under Scrutiny. Paris.

School Evaluation Guidebook 2010-2016. Seoul: Korean Educational Development Institute (KEDI).

Scriven, M. (1967). The methodology of evaluation. In R. E. Stake (Ed.), *AERA monograph series on curriculum evaluation*, No. 1. Chicago: Rand McNally.

Scriven, M. (1991). *Evaluation thesaurus* (4th ed.). Thousand Oaks, CA: Sage.

Sonnichsen, R. C. (2000). *High impact internal evaluation*. Thousand Oaks, CA: Sage.

Stufflebeam, D. L. (1969). Evaluation as enlightenment for decision making. In W. H. Beatty (Ed.), *Improving educational assessment and an inventory for measures of affective behavior*. Washington, D. C.: National Education Association.

Stufflebeam, D. L., Foley, W. J., Gephart, W. J., Guba, E. G., Hammond, R. L., Merriman, H. O., & Provus, M. M. (1971). *Educational evaluation and decision making*. Itasca IL: Peacock.

Stufflebeam, D. L. (1974). *Meta-evaluation* (Occasional Paper No.3). Kalamazoo: Western Michigan University, December.

Stufflebeam, D. L., & Chris L. S. Coryn (2014). *Evaluation Theory, Models, and Applications*. Second Edition, Jossey-Bass.

Tyler, R. W. (1950). *Basic Principles of Curriculum and Instruction*. Chicago, Ill.: University of Chicago Press.

Werner, W. (1978). Evaluation: Sense-making in school programs. In T. Aoki (Ed.), *Curriculum evaluation in a new key*. (Monograph Series No. 1). Vancouver: University of British Columbia, Center for the Study of Curriculum and Instruction, Faculty of Education. (ED 153 907).

(Accepted 17th September, 2017)

【Research Note】

Evaluation of Elementary and Junior High Schools in Taiwan: Practices and Challenges

Shu-Huei Cheng
National Taiwan Normal University
shcheng@ntnu.edu.tw

Abstract

This study analyzed the characteristics of elementary and junior high school evaluation in Taiwan. Document analysis was used to examine the evaluation systems of seven local governments across Taiwan. The context of Taiwanese educational evaluation was first analyzed. Then, referring to a theoretical framework, this study investigated the evaluation practices of elementary and junior high schools in terms of purposes, organization, evaluators, procedures, criteria, methods, reporting, and intended evaluation use. Finally, challenges involving school evaluation systems were discussed. This study highlighted how government-mandated school evaluation practices are structured and challenged in the context of education reform.

Keywords

school evaluation, elementary school, junior high school, Taiwan

Public demand for quality education has intensified in the context of global competitiveness. Education systems are expected to produce educated workforces that can adapt to rapid and unexpected changes in a knowledge-based economy (OECD, 2013). School evaluation is an effective approach that provides information for educational improvement and accountability (Nevo, 2009).

Numerous countries have developed their own school evaluation practices; nevertheless, the question of how to build a sound evaluation system remains challenging (Faubert, 2009). This study used a Taiwanese case to analyze school evaluation practices for two reasons. First, the Taiwanese educational system shares many features with other East Asian societies. For example, education is highly valued, and its governance tends to be centralized (Peng & Lee, 2009). Second, Taiwanese evaluation cases have rarely been investigated in non-Chinese-language publications. Because of Taiwan's similar cultural and governance contexts, a Taiwanese case may provide insight into school evaluation systems, challenges, and potential solutions for other East Asian societies.

Therefore, this study used document analysis to investigate current practices in elementary and junior high school evaluation in seven cities and counties across Taiwan. This paper begins by illustrating the context of educational evaluation in Taiwan, followed by the analytical framework and methodology. The research findings, in

terms of the key features of school evaluation, are subsequently analyzed, and the results and conclusions are presented.

1. Context of Educational Evaluation in Taiwan

Similar to numerous other East Asian countries (Peng & Lee, 2009), governments are the key bodies that commission educational evaluations in Taiwan. The first educational evaluation was initiated in the 1960s because of the Taiwanese government's involvement in a United Nations program. Following this externally supported evaluation, educational evaluation was implemented intermittently at the kindergarten, elementary, junior high school, senior high school, and university levels. Not until the 1990s did educational evaluation expand and gradually institutionalize (Guo, 2000). In a context of economic growth, deregulation, and decentralization, the central government transferred educational authority to local governments and schools. School-based management has been promoted and decision-making within schools is shared with teachers and parents (Pan & Chen, 2011).

Although schools are deregulated and decentralized, under the Education Basic Law, evaluation has been advocated as a policy tool to ensure educational quality. Evaluating kindergartens, senior high schools, and universities is required by law. Attending elementary and junior high schools is compulsory for students aged 6–12 years and 13–15 years, respectively. According to the Elementary and Junior High School Act, local governments must evaluate principals to establish a reference to be used when considering reappointments for principals. Some local governments conduct school evaluation to more fully understand the functions and performance of schools as well as to serve as an alternative to principal evaluation (Lin & Wang, 2017). The practices of school evaluation differ across the 22 cities and counties of Taiwan. Cheng (2007) and Hsieh (2011), investigating school evaluation practices in 2005 and 2009, respectively, have reported that half of local governments have developed school evaluation systems. In light of policy changes over time, this study analyzed the contemporary variations in school evaluation systems among the sampled Taiwanese cities and counties.

2. Analytical Framework and Methodology

Reviews of educational or school evaluations have mainly focused on systematic investigation and judgment of the value of the evaluand. Sanders and Davidson (2003: 807) defined school evaluation as the "systematic investigation of the quality of a school and how well it is serving the needs of its community." The process involves collecting systematic data and applying defensible criteria related to the nature and quality of educational evaluands (Fitzpatrick, Sanders, & Worthen, 2011; Nevo, 1995; Owen, 2007). School evaluation is essential to school development because it provides a more profound understanding of school practices, directions for improvement, references for decision-making, and records for accountability (Sanders & Davidson, 2003; Stufflebeam, 2003).

School evaluation is a complex system, and its practices vary among countries. Faubert (2009) analyzed the key features of school evaluation practices in OECD countries in terms of purposes, scope, users, agencies, procedures, and use. The evaluation procedures can involve planning, implementation, reporting, and use (Fitzpatrick et al., 2011; Owen, 2007). In this study, school evaluation was defined as a systematic process for assessing school quality on the basis of data collection, analysis, and criteria application. By using the evaluation results, stakeholders in school evaluation can enhance their understanding, decision-making, action for improvement, and responsiveness to accountability demands.

Referring to the analytical frameworks, this study used document analysis to investigate the current practices of elementary and junior high school evaluation in Taiwan. School evaluation official documents for the 2016 (after August 1) to 2017 (before July 31) academic year were included in the analysis. Evaluation documents covering the previous academic years and discussions on school evaluation on the internet were used to supplement understanding of school evaluation practices.

To locate the documents that were required for this study, I used local government websites to seek documents related to school evaluation. Furthermore, I searched Google for "school evaluation" and the names of cities and counties in Taiwan to explore other related information. In addition to an online search, I was granted access to a city's school evaluation plan, which cannot be accessed on the internet. The document search continued until July 17, 2017. To ensure accuracy, I triangulated different document search results.

This study ultimately included nine evaluation systems from seven cities and counties of Taiwan, namely Taipei City, New Taipei City, Hsinchu County, Taichung City, Yunlin County, Kaohsiung City, and Taitung County. Notably, Taipei City and Kaohsiung City governments have different evaluation systems for elementary and junior high schools. The school evaluation plans for all nine systems, publicly available evaluation results, and training handouts were included in this study. Content analysis was performed to categorize the documents and measure the frequency of categories on the basis of the analytical framework.

3. Findings

This study analyzed evaluation practices in elementary and junior high schools in Taiwan in terms of purposes, organization, evaluators, procedures, criteria, methods, reporting, and intended use.

3.1. Purposes

Analysis of evaluation documents indicated that school evaluation systems have many different purposes. All of the school evaluations in this study are used to provide an understanding of the overall situation, strengths, policy implementation, problems encountered, and the extent to which schools serve their students. This understanding provides schools with a direction for improvement and provides the local government with references for making decisions to promote educational quality.

Additionally, some of the school evaluation systems are used to demonstrate school performance and exhibit accountability. For example, two local governments use school evaluation results as a reference for deciding on principal reappointments. One city government makes school performance information publicly available.

Another purpose of school evaluation, which is rarely discussed in the Western literature, is to integrate multiple school-level evaluations. Local governments in Taiwan tend to conduct evaluation of recently implemented policies and programs. For example, one local government uses a school evaluation system that includes criteria related to gender equity in education, physical and health education, integrating information technology into instruction, campus accessible facilities, disaster prevention education, campus safety programs, and teacher mentoring programs. This local government uses school evaluations to collect information on these seven criteria instead of undertaking separate evaluations for each one.

3.2. Organizations and Evaluators

The school evaluation systems that were included in this study all integrate two levels of organization into school evaluation. The first level is governmental. Local governments establish committees that are responsible for

designing, coordinating, supervising, and implementing evaluations as well as meta-evaluations. Depending on the local government, the members of the committees can include scholars and key stakeholders, such as school administrators, teachers, parents, and officials from related divisions of the local government. Three of the city governments contract a major part of school evaluation tasks out to universities.

The evaluators, who are responsible for on-site evaluation, are appointed by the local governments according to their expertise in aspects of school evaluation. All of the school evaluation systems in this study included scholars and school practitioners as evaluators. The practitioners may include excellent retired principals and current teachers within or outside the city or county, depending on the regulations. One local government responds to parental concerns and invites parent representatives as evaluation participants to collect data on special education, but does not require them to write evaluation reports.

The second level of organization is the school level. Local governments require that every school form a committee responsible for self-evaluation. Chaired by the principal, the committee must comprise school administrators, teachers, and parents. Two local governments specify the number of committee members—five to nine and at least six respectively. These two local governments encourage schools to invite scholars to join the self-evaluation committee.

3.3. Procedures

Reviews of school evaluation documents indicated that every school is externally evaluated every 2, 3, or 4 years depending on the local governments' requirements. After analyzing the evaluation procedure, this study determined that school evaluations have six tasks in common. The first task is to hold a seminar for school personnel, including the principal, school administrators, and teachers. This evaluation seminar is held and paid for by the local government, which determines the content of the seminar.

The seminar for school personnel is held to explain the purpose, scope, timing of evaluation, and the evaluation methods. Take one evaluation system detailed at a seminar for example. To prepare schools for the implementation of a new evaluation system, this city government conducted a 5-day seminar on self-evaluation report writing and a 1-day seminar on teacher observation, a method required for school evaluation. The instructors included university scholars and experienced principals.

The second task of the procedure is self-evaluation. Each school is required to conduct a self-evaluation and write an evaluation report. As mentioned before, a self-evaluation committee is formed that includes key stakeholders who review the school's situation on the basis of the requested criteria and format.

The third task of the evaluation procedure is to hold a seminar for the government-appointed evaluators. This seminar aims to assist evaluators in understanding the rationale of school evaluation, the ethical principles involved, and report writing before the evaluation. For example, one city government held an 18-hour seminar to prepare evaluators for a new evaluation system. Notably, several evaluation documents detail ethical principles for on-site evaluators, such as recusal, confidentiality, and commitment to evaluation.

The fourth task of the procedure is on-site evaluation, a major phase in the data collection that is used to develop evaluation reports. All of the evaluation systems require that evaluators use multiple methods to ensure data accuracy. In addition to a school's presentation, observations, interviews, and document analysis are required in order to provide qualitative and quantitative data. An on-site evaluation takes a half-day or 1 day to conduct. An exception is one local government's 2-day on-site evaluation. Another evaluation system allows flexibility to provide for 1 to 2 days of on-site evaluation depending on school size. The length of time arranged for on-site evaluation affects the number of observations and interviews that can be organized.

The fifth task of the procedure is the examination of the evaluation results and the appeals process. After the

evaluation reports are finalized, local governments convene meetings to confirm the results. After receiving their evaluation reports, schools have the right to appeal in certain circumstances, such as when the evaluation result is harmful because it is based on data that contradicts the school's real circumstances. The meeting finalizes the evaluation results after consideration of the appeal and the collected data. The documents included in analysis, however, do not identify the evaluators who finalize the results. In the author's experience with two of the evaluation systems, this meeting mainly includes the representatives of the on-site evaluators.

The sixth task of the procedure is follow-up, which is specific to low-performing schools. A self-improvement plan is required to be implemented on the basis of the suggestions provided by the evaluation report. Supervision and further evaluation may be conducted depending on the grades or ranks that a school receives in its evaluation.

Besides the aforementioned six tasks, one local government explicitly includes an evaluation design stage, which involves focus groups and a public hearing before the school seminar stage. Two local governments include meta-evaluation to enhance quality of evaluation. Additionally, one of the local governments arranges a seminar in which high-performing schools that were identified based on evaluation results share their practices with other schools. According to Lin (2004) and Wu (2002) , sharing best practices is one of the approaches that are acknowledged to strengthen the evaluation use.

3.4. Criteria

Evaluation criteria provide a basis for judging school quality (Fitzpatrick et al., 2011). A review of evaluation documentation revealed that local governments use multiple criteria that can be categorized into five to ten dimensions. Four of the evaluation systems adopt five dimensions from the criteria, and two systems adopt six dimensions. The other three systems use seven, nine, and ten dimensions from the criteria, respectively.

Although the local governments use different categories, they generally cover the vital functions and outcomes of schools, such as leadership, management, curricula, instruction, professional development, student affairs, counseling and special education, physical environment, equipment and resource utilization, public relationship and parental involvement, student learning, as well as school growth and features. Some of the local governments also incorporate the national supervision criteria and legal requirements. In addition to public schools, some of the school evaluation systems include private schools and develop criteria related to the school boards, whose role is to direct private-school development. Additional analysis revealed that the evaluation systems that adopt five or six dimensions combine several criterion categories, such as curricula, instruction, and professional development, as well as leadership and management.

To properly respond to the concerns of international and Taiwanese educators, student learning, principal leadership, and school features are detailed in the following paragraphs. Student learning is stressed in school evaluations worldwide. Among the evaluation systems I studied, four in particular categorize student learning as a dimension or subdimension of the evaluation criteria. The criteria signify the values of whole-person education and include students' character and cognitive and noncognitive learning. These criteria related to learning focus on how well schools enhance student learning, and to what extent students are involved in learning activities. The criteria also cover to what extent students demonstrate their growth and learning outcomes.

Principal leadership is essential to school quality and is usually a key dimension or subdimension of evaluation criteria in Taiwanese school evaluation. However, the inclusion of principal leadership in school evaluation causes confusion over the differences between school evaluation and principal evaluation (Lin & Wang, 2017; Tang, Chen, Kuo, & Chu, 2012). In response to the confusion, one city government removed principal leadership to a supplementary dimension of the evaluation criteria. Thus, the data on principal leadership are still collected and analyzed during school evaluations, but it is presented separately from the overall school evaluation results.

School evaluation systems typically use the same quantitative criteria to judge school quality, regardless of differences in school size, resources, and community and student characteristics. In addition to the preplanned evaluation criteria applied to all schools, all of the evaluation systems in this study include qualitative criteria that capture the features of individual schools.

3.5. Methods

Evaluation methods are used to collect data relevant to the evaluation criteria. Theoretically, data collection relies on self-evaluation and on-site evaluation. However, the school evaluation documentation examined in this study includes on-site evaluation rather than self-evaluation methodology, indicating that on-site evaluation is expected to be more rigorous than self-evaluation.

Multiple methods are preferred in all of the school evaluation systems in this study. Among the evaluation methods, observations, interviews, and document analyses are frequently used. Observations used in school evaluation focus on physical environments and school instructional and learning activities. Some school evaluation systems prioritize classroom observations conducted by excellent teachers with observation forms. The classroom observations are analyzed in terms of the criteria of instruction and are also usually provided for individual teachers' reference.

School administrators and teachers are interviewed in all of the studied evaluation systems. Some include interviews with students, parents, and community members to capture key stakeholders' viewpoints regarding the schools. Parents and community members are sampled by the schools, and the other interviewees are sampled on-site by the evaluators in accordance with the criteria established in the evaluation systems.

Self-evaluation reports and documents prepared by the schools, such as plans and records, are also reviewed on the basis of the evaluation criteria. In particular, four local governments provide evaluators with access to digital documents for review before on-site evaluation.

Only one local government in this study uses online surveys for school evaluation. Respondents are randomly recruited from among the teachers, administrators, parents, and students, which can result in a larger sample than is possible in interviews. The survey contains both Likert-scale and open-ended questions, which reflect the evaluation criteria. In addition to the qualitative comments, the evaluation team provides the evaluators with the means, percentages, and cross-analysis of demographics and responses before on-site evaluation. The means of similar-sized schools and other schools in the same city are also provided for comparison.

3.6. Reporting and Intended Evaluation Use

Reporting and evaluation uses are crucial because the evaluation results should be communicated with key stakeholders to engender positive effects on educational practices (Owen, 2007). Document analysis indicated that the reports are used to inform schools and local governments. Although the format varies, the reports usually include the evaluators' appraisal, such as grade and rank, and qualitative descriptions of the schools' strengths and weaknesses along with suggestions for the schools and government. One city government also makes evaluation results publicly available on the Internet. The information provided includes the features of all evaluated schools and the names of the schools that receive the top two out of five ranks.

The evaluation documentation describes the follow-up procedures. For example, schools must write an improvement plan and act in response to suggestions. Local governments adopt supervision or evaluation measures for schools with unsatisfactory evaluation results. The documentation indicates that the evaluation process includes provision for prizes to be awarded to schools and their personnel to recognize high-level efforts and performance. Furthermore, two local government evaluation systems consider evaluation results when awarding another prize that

acknowledges excellent or featured schools.

4. Discussion and Conclusion

School evaluation is acknowledged as essential to educational improvement and accountability demands. However, developing a sound school evaluation system remains challenging. This study used the Taiwanese elementary and junior high school evaluation systems to provide topics for reflection in terms of purposes, organization, evaluators, procedures, criteria, methods, reporting, and intended evaluation use.

School evaluations in Taiwan are primarily government mandated and have become increasingly systematic and common since the 1990s. Consistent with the literature and the evaluation practices of OECD countries, the major purposes of school evaluations are improvement, decision-making, and accountability (Faubert, 2009; Nevo, 2006). School personnel face pressure from frequent school monitoring, on-site visits, and external evaluations. One additional school evaluation purpose in Taiwan, is the integration of multiple school-level evaluations by serving the information needs of other evaluation-related activities. Nevertheless, it remains a challenge because different evaluations may have unique information needs depending on the circumstances (Cheng, 2016).

The school evaluation systems across local governments have developed six tasks in common, including short-term seminars for school personnel and external evaluators, self-evaluation followed by an on-site external evaluation, examination of evaluation results, and a follow-up to evaluation. School evaluations are organized by both governments and schools. Local governments are responsible for designing and conducting the main school evaluation; schools are required to assemble committees to conduct self-evaluations. The two-level composition of the evaluation groups is designed to include scholars and various stakeholders in an effective manner. Their involvement can enhance their learning from evaluation process and findings (Johnson, Greenseid, Toal, King, Lawrenz, & Volkov, 2009).

Although the school evaluation systems included in this study rely heavily on external evaluation, self-evaluation is necessary to comply with the external evaluation process, and schools are encouraged to conduct it as part of school management. Nevertheless, Cheng and King (2017) reported that self-evaluation capacity in Taiwanese elementary and junior high schools should be enhanced in terms of evaluation culture, evaluation infrastructure, and human resources, and governments and schools must overcome any challenges that hinder effective self-evaluation.

External evaluators, including scholars, principals, and teachers, are appointed for their expertise in education rather than their evaluation competence. To develop relevant evaluation skills, local governments usually hold short-term seminars for evaluators. The evaluation systems also require evaluators to make judgments by adhering to ethical principles, using multiple data sources, and holding discussions with other evaluators. The empirical evidence indicates that evaluation audiences continue to expect evaluators to achieve accurate results and feasible suggestions (Cheng, 2016; Huang, 2012).

Consistent with the literature, school evaluation in Taiwan uses multiple criteria and methods to assess schools' major processes and outcomes (Nevo, 1995; Stufflebeam, 2003). In accordance with the goals of educational reform, student learning is emphasized in school evaluation. School evaluation systems in Taiwan focus heavily on how schools address student learning needs. The criteria related to student learning include the learning outcomes for students' character and cognitive and noncognitive skill development, reflecting the value of whole-person education rather than overreliance on academic achievement.

Although multiple criteria and methods are useful for assessing the many aspects of school quality, this complexity may increase the burden on schools when they prepare their documents to write the self-evaluation

reports (Cho, 2007), and may also pose challenges for half-day or 1-day on-site evaluation (Cheng, 2007).

The evaluation documentation outlines one attempt to mitigate the constraints of the time limit for on-site evaluation by providing evaluators with digital documents and online survey results to be analyzed beforehand. Another method attempts to reduce the data collection burden on schools by effectively using other existing data sources. Nevertheless, some school personnel doubt whether the evaluators can completely characterize their schools through their limited time on-site despite their use of multiple methods (Cheng, 2007; Lin, 2014). The correct balance between assessing the complexity of school quality and simplifying data collection is challenging. The burden caused by school evaluation is still a concern.

Finally, evaluation reports are usually presented to the school being evaluated and relevant government divisions. One local government in this study made evaluation results publicly available, a practice that is controversial because of its potential drawbacks (Faubert, 2009). The evaluation results shown on the government's website included qualitative descriptions of the schools' features and the names of the schools that received satisfactory evaluation results. Although the names of schools that received unsatisfactory evaluation results were not posted online, follow-up to evaluations were conducted, comprising improvement plans, supervision, and additional evaluations. This analysis of reporting and evaluation use indicated that Taiwanese school evaluation emphasizes schools' self-improvement and the role of local governments in supervision and assistance. Market-type accountability is not a fundamental aim of Taiwanese school evaluation. Concerning school improvement, governments and schools still need to overcome any challenge that hinders evaluation use of process and findings (Cheng, 2016; Huang, 2012).

Aside from summarizing the evaluation process, the evaluation report is intended to acknowledge best practices and schools' efforts by awarding prizes to schools and personnel that receive positive evaluation results. School personnel acknowledge the positive effect that awarding prizes has on morale (Chen, 2002); nevertheless, they generally hope that the number of school evaluations is reduced mainly because of the fatigue attributable to preparing for evaluation (Li & Wen, 2016).

This study analyzed the characteristics of elementary and junior high school evaluation practices in Taiwan and the challenges that have been encountered, by using document analysis and including the evaluation practices of the 2016–17 academic year. Although the accuracy and completeness of the documents were carefully examined, a gap in the evaluation design and implementation may exist. Future studies should conduct case studies involving multiple methods to investigate the up-to-date implementation and effects of school evaluation practices.

References

Chen, C. H. (2002). *Research on school evaluation of junior high and elementary schools in Taitung County* (Unpulished master thesis). National Taitung University, Taitung, Taiwan.

Cheng, C. C. (2007). *Research on evaluation criteria and implementation in elementary and junior high schools*. Taipei, Taiwan: Psychological.

Cheng, S. H. (2016). School evaluation use among local educational agencies. *Contemporary Educational Research Quarterly*, 24 (4), 33-69.

Cheng, S. H., & King, J. A. (2017). Exploring organizational evaluation capacity and evaluation capacity building: A Delphi study of Taiwanese elementary and junior high schools. *American Journal of Evaluation, 38* (4), 521-539.

Cho, I. P. (2007). *A study on self-evaluation of school evaluation in Taipei elementary schools* (Unpulished master thesis). University of Taipei, Taipei, Taiwan.

Faubert, V. (2009). *School evaluation: Current practices in OECD countries and a literature review*. Retrieved from http://dx.doi.

org/10.1787/218816547156

Fitzpatrick, J. L., Sanders, J. R., & Worthen, B. R. (2011). *Program evaluation: Alternative approaches and practical guidelines* (4th ed.). Boston, MA: Pearson.

Guo, C. Y. (2000). *School-based evaluation*. Taipei, Taiwan: Wu-Nan.

Hsieh, H. C. (2011). *A study on implementation of school evaluation system for the junior high and elementary schools in Taiwan* (Unpublished master thesis). National Hsinchu University of Education, Hsinchu, Taiwan.

Huang, C. H. (2012). A research on the factors that influence the uses of curriculum and instruction evaluation findings in elementary schools. *Curriculum & Instruction Quarterly, 15* (3), 1-26.

Johnson, K., Greenseid, L. O., Toal, S. A., King, J. A., Lawrenz, F., & Volkov, B. (2009). Research on evaluation use: A review of the empirical literature from 1986 to 2005. *American Journal of Evaluation, 30* (3), 377-410.

Li, A. M., & Wen, Y. S. (2016). Difficulties and challenges of educational reform: Problems and strategies of school institutional evaluation. *School Administrators, 101*, 1-14.

Lin, S. C. (2014). *The study of elementary school evaluation in Taichung city* (Unpublished master thesis). Providence University, Taichung, Taiwan.

Lin, W. J., & Wang, L. Y. (2017, October). *The problems of combining principal evaluation and school evaluation*. Paper presented in the annual conference of Center for Educational Research and Evaluation: International Conference on Research Informed Policy and Practice, Taipei, Taiwan.

Nevo, D. (1995). *School-based evaluation: A dialogue for school development*. Oxford, OX, UK: Pergamon.

Nevo, D. (2006). Evaluation in education. In I. F. Shaw, J. C. Greene, & M. M. Mark (Eds.), *Handbook of evaluation: Policies, Programs, and practices* (pp. 441-460). Thousand Oaks, CA: Sage.

Nevo, D. (2009). Accountability and capacity building: Can they live together? In K. E. Ryan & J. B. Cousins (Eds.). *The Sage international handbook of educational evaluation* (pp. 291-303). Thousand Oaks, CA: Sage.

OECD (2013). Education today 2013. Retrieved from http://abdigm.meb.gov.tr /meb_iys_dosyalar/2013_12/20023619_educationtoday2013theoecdperspective.pdf

Owen, J. M. (2007). *Program evaluation: Forms and approaches (3rd ed.)*. New York, NY: Guilford.

Pan, H. L.W., & Chen, P. (2011). Challenges and research agenda of school leadership in Taiwan. *School Leadership & Management, 31* (4), 339-353.

Peng, S. S., & Lee, J. C. K. (2009). Introduction. In S. S. Peng & J. C. K. Lee (Eds.), *Educational evaluation in East Asia: Emerging issues and challenges* (pp. 1-10). New York, NY: Nova.

Sanders, J. R., & Davidson, E. J. (2003). A model for school evaluation. In T. Kellaghan, D. L. Stufflebeam, & L. A. Wingate (Eds.), *International handbook of educational evaluation: Part Two: Practice* (pp.807-826). Dordrecht, The Netherlands: Kluwer.

Stufflebeam, D. L. (2003). Institutionalizing evaluation in schools. In T. Kellaghan, D. L. Stufflebeam, & L. A. Wingate (Eds.), *International handbook of educational evaluation: Part Two: Practice* (pp.775-805). Dordrecht, The Netherlands: Kluwer.

Tang, Z. M., Chen, M. J., Kuo, C. Y., & Chu, J. R. (2012). A preliminary study on the dimensions of principal evaluation in elementary and junior high schools. *Journal of Education Research, 224*, 5-17.

(Accepted 17th September, 2017)

【Research Note】

Note on the Characteristics of Japanese School Evaluation: Seven factors obstructing successful implementation

Akihiko Hashimoto
National Institute for Educational Policy Research, Japan
aki@nier.go.jp

Abstract

This note describes the current trend of school evaluation in Japan, points out its characteristics, and considers what is wanted. Starting with a broad overview of the system of school evaluation, this note will introduce the idea that school evaluation has functioned more poorly than expected, and the note will describe seven specific factors that may have obstructed the successful implementation of school evaluation. These seven factors are: 1) a lack of accurate understanding of the meaning of the Japanese word "evaluation"; 2) school evaluation not linked to schools' goals; 3) poor indicators for evaluation; 4) evaluation results not utilized for school improvement; 5) an absence of appropriate evaluators or coordinators; 6) a lack of training courses; and 7) poor funding. These issues do not appear in all Japanese schools, and every factor requires further investigation; however, this note provides the hypothetical characteristics of typical Japanese schools. Researchers may be able to share tentative conclusions for comparison between school evaluation systems as they are implemented in other countries.

Keywords

comparative study, evaluation study, implementation, Japan, school evaluation

Introduction

In Japan, the phrase "school evaluation" is understood without question even though many people may not know much about school evaluation. School evaluation has been introduced and established as part of the national administration of education in almost all schools. However, as later sections of this paper will demonstrate, there are criticisms that school evaluation has not been implemented in an efficient way.

The first section of this note will outline the process of the implementation of school evaluation in Japan over the past two decades. It will deal with the processes of school evaluation, including evaluation procedures, indicators, contents, methodology, and problems. The second section will analyze the seven typical features of Japanese school evaluation. This note does not intend to prove that these features apply at all times, but may present hypothetical threads to be investigated. Through this note, researchers may be able to share tentative conclusions for comparison

between school evaluation systems as they are implemented in other countries, and to highlight the necessity for further studies of each country's process of school evaluation.

1. Brief history and outline of the Japanese school evaluation system

1.1. The legislation for and purpose of the school evaluation system

Since the 1950s, many trials and approaches to school evaluation were attempted throughout the country but school evaluation did not spread nationwide (Kioka 2005). It was only after the 1990s that school evaluation became a national system. After the creation of the Administrative Reform Council in November 1996, under the leadership of then-Prime Minister Ryutaro Hashimoto, the government aimed to introduce the administrative practices of the New Public Management. Implementing school evaluation was considered one of the main pillars of the Hashimoto administration's education reforms. In response to these political trends, the Ministry of Education, Science and Culture's Central Council for Education *Report on How Local Administration on Education Should Be*, issued in September 1998, recommended the implementation of various evaluation techniques in educational administration and school operation. The report noted also that self-evaluation must be carried out to determine how the school's goals are met and the kinds of plans and programs that are in action to enable parents and community members to engage with the school (Ministry of Education, Science and Culture 1998).

Complex controversies over school evaluation in governmental councils were related to educational reforms. On the one hand, council members representing the business world claimed that the school evaluation system must operate in a competitive context, through the introduction of external evaluation; these members claimed that school evaluation should provide parents with the information needed to make good choices about their children's schools, or it should deliver basis for economic distribution to schools. On the other hand, members with backgrounds in education field insisted that evaluation should not be used to rank schools or to discriminate against certain schools. The Ministry of Education, Culture, Sports and Scientific Technology (hereafter abbreviated as MEXT[1]) eventually balanced these arguments to set up an evaluation system to empower schools toward improvement (Hirota and Ikeda 2009).

Mandating that schools implement school evaluation was another contentious issue in the process. In March 2002, the "Standard for the Elementary School Establishment" and other similar MEXT ministerial ordinances were modified, and each school was required to "make efforts" to evaluate its activities and disclose the results of the evaluation to the public.[2] The Council for Regulatory Reform (2003) proposed promoting schools' self-evaluation through methods including making evaluation mandatory through a ministerial ordinance.[3] The proposal to make school evaluation mandatory also appears many times in other council reports. MEXT gradually determined its plan of action when the Central Council for Education issued a report titled *Creating a New Era of Compulsory Education* on October 26, 2005. In this report (2005), the Council proposed that school evaluation guidelines with clearly outlined principles should be prepared for the reference of schools and local municipalities, and that self-evaluation at each school should be made mandatory. In terms of external evaluation, the report only mentioned the importance of further examination, including third-party evaluation and national involvement.

In June 2007, the "School Education Law" was amended and, in Article 42, school evaluation was made a duty for all schools, with a provision ordering information providing by schools in Article 43. A translation of Article 42 is shown as Table 1.[4] It is important to note that the purpose of school evaluation is stipulated as "to improve school operation" or "to improve the level" of schools' education.

Article 43, promulgated with Article 42, required schools to provide persons related to the school, such as parents and community members, with information about school operation in order to foster partnerships between the

Table 1　School Evaluation in the Law, 2007

School Education Law Article 42

Based on the regulations stipulated by the Minister of Education, Culture, Sports and Scientific Technology, elementary schools shall endeavor to improve the level of their education by carrying out evaluations of the conditions of their educational and other activities; then, in accordance with the results, implement necessary measures to improve school operations.

Source: Author's translation of the School Education Law of Japan, Article 42

school and parents and community members. This provision was intended to assure that accountability could be added as a purpose of school evaluation, in addition to school improvement.

In October 2007, MEXT promulgated the "Enforcement Regulations of the School Education Law," which details two types of school evaluation.[5] These are, first, "self-evaluation" carried out by school staff and, second, a kind of external evaluation "done by parents and other persons related to the school following the results of self-evaluation," as described in articles 66 and 67, respectively. Self-evaluation is made mandatory, as are requirements to make efforts to implement the external evaluation. Article 68 requests that the results of both kinds of evaluation be reported to the founder of the school which is, in many cases, the board of education.

Thus, the legal system for school evaluation was developed. The purpose of the school evaluation system was clearly stated as for schools "to improve the level" of their education by implementing "necessary measures to improve school operations" following the results of the evaluation.

1.2. School Evaluation Guidelines

School evaluation guidelines were first mentioned on June 1, 2005 at a meeting of the Council on Economic and Fiscal Policy chaired by then-Prime Minister Junichiro Koizumi, where members proposed, "guidelines should be formulated during fiscal 2005 to support the conduct of schools' external evaluation and the reporting of results." After this proposal, the then-Minister of Education, Culture, Sports, Science and Technology, Akira Nakayama, explained the Ministry's plan noting that the ministry "will make the self-evaluation and the announcements of its results at each school obligatory, and will work to promote the external evaluation carried out by parents and community members. The proposed guidelines should also be actively promoted."[6] The Cabinet soon brought the idea to realization. On June 21, *Basic Policies for Economic and Fiscal Management and Structural Reform 2005* was endorsed by the Cabinet, who declared, "guidelines for the conduct and reporting of schools' external evaluation should be formulated during fiscal 2005" (Cabinet Office 2005).

Debates in the field of education followed this trend. In the Central Council for Education's *Creating a New Era of Compulsory Education*, issued on October 26, 2005, the Council proposed that, as mentioned above, school evaluation guidelines should be prepared and self-evaluation at each school should be made obligatory. Though external evaluation was not concretely mentioned in this report, the "necessity of investigating further measures of improving examinations, including third-party evaluation" was noted.

In addition to arguments made at these committees, school evaluation guidelines were prepared under MEXT's initiative. Following national basic policy, on August 10, 2005 MEXT appointed a School Evaluation System Study Group to discuss and endorse the draft guidelines prepared by the Ministry.[7] At the Study Group's first meeting on February 20, and at the second (and last) meeting on March 20, 2006, comments were put to the draft guidelines. As per the published summary of the meetings' proceedings, the main comments are summarized in Table 2.[8]

Table 2 Main comments on the Draft School Evaluation Guidelines from the School Evaluation System Study Group

1st meeting: February 20, 2006

Important to enrich school management performance.
Important to specify school goals.
Important to regard the evaluation targets and indicators as written in the guidelines as examples only and to form original targets and indicators, as suitable for each school. It may be also appropriate for school founders to formulate common targets and indicators for the whole district.
Better to allow each school to choose the timing for the collection of the results of self-evaluation, and that this collection not happen only at the end of fiscal year.
Necessary to enrich professional training to increase the knowledge etc. of the leading teachers and external-evaluators at each school so that school evaluation will be carried out properly.

2nd meeting on March 20, 2006

Necessary to set up an evaluation group within the school to manage self-evaluation.
Better to leave external-evaluation flexible, as external-evaluation is difficult to conduct annually.
Important to collect school information by performing self-evaluations.
Better to declare in the "foreword" that the *Guidelines* are not binding schools in terms of evaluation methods.
Better to mention "class management" among the evaluation indicators.
Important to mention and emphasize that the school's founder shall develop conditions to support and improve the school after the school evaluation results.
Better to specify the incentives of conducting school evaluations: Giving the power for taking action, giving support after evaluation results, and dispatching teacher consultants as advisors, according to the evaluation results, are some examples of incentives.
Necessary to respect schools' input as well as students' performance indicators.

Source: Author's translation, from the MEXT web site.

As Hirota and Ikeda (2009) point out, viewing the outline of the proceedings, it appears that the study group's comments mostly support MEXT's arguments on school evaluation. It is remarkable that the terms seen in the Study Group's web page are then put into the "Foreword" section of the final guidelines, which will be explained below.

The *Guidelines for School Evaluation at Compulsory Education Stage* (hereafter referred to as the *Guidelines*) were first issued in March 2006 and were revised three times by March 2016. The *Guidelines* show samples of school-evaluations, in terms of "purposes," "targets," "techniques," "viewpoints," "indicators," "reporting," "result utilizing" and so on. Local boards of education or local schools design their own school evaluation system according to the *Guidelines* and school conditions.

By 2007, MEXT was prepared to begin a mandatory school evaluation system. In June 2007, as noted above, the *School Education Law* was amended to make school self-evaluation aimed at the improvement of educational activity its main component. Though mandatory, the system was accompanied by national guidelines that are not binding but that leave space for discretion on the part of schools and local authorities, in terms of choosing specific methods and procedures for school evaluation.

1.3. The characteristics and structure of the school evaluation system

The last step to building the national structure of the school evaluation system was clarifying the boundaries and relationships between self-evaluation and external evaluation. In the 2006 *Guidelines*, school evaluation was simply divided into "self-evaluation" and "external evaluation by persons related to the school, guardians, and community residents." The characteristics of external evaluation were not yet specified, particularly in terms of the function and position of the so-called "third-party evaluation."

A Panel of Research Experts for Promoting School Evaluation was appointed in July 2006. The panel mainly consisted of the previous members of The School Evaluation System Study Group, which had commented on first draft Guidelines. The panel's task was to investigate the measures necessary for the promotion of school evaluation in terms of "planning improvements to the operation of schools." To aim to plan these "improvements" was an outcome of following the school evaluation characteristics stipulated in the amended School Education Law of 2007.

The Panel of Experts argued that "third-party" should refer to specialists who are staff members of universities or educational research institutes, or other scholars. It is notable that educational specialists or researchers could take a leading position in such a group.[9] The Panel concluded that a "third-party evaluation" must be distinguished from an external evaluation "by persons related to the school." While an evaluation carried out by parents and community residents may accomplish citizen participation, third-party evaluation contributes to schools through its professionalism. The Panel highlights "professionalism" from both pedagogical and managerial viewpoints, which aim to provide schools with lacking skills and resources. The Panel also focuses on "objectiveness" and "independence," noting that evaluators' comments and proposals should be fresh and fair, affecting many people concerned with spontaneous changes to school activities. In sum, the Panel proposed third-party school evaluation as "a professional and objective evaluation to elevate the quality of school operation," as carried out by "an independent evaluation organization."[10]

The *Guidelines* were revised in January 2008. At this time, school evaluation in upper secondary schools (high schools) was considered, and the guidelines' title was revised to *School Evaluation Guidelines*. The word "external evaluation" was no longer used and was replaced by the new category of "evaluation by persons related to the school."[11] In addition to this "evaluation by persons related to the school," the *Guidelines* clarified the role of "third-party evaluation" describing it as "a professional and objective (third-party perspective) evaluation done by specialists who do not have a direct relationship with the school." While the guidelines did not discuss third-party evaluation methods in detail, the framework for Japanese school evaluation was established by the January 2008 revision.

Thus, through the law, regulations and the *Guidelines*, the central government had proposed a three-tier model of school evaluation, as shown as Table 3. Through the 2007 law, the purpose of school evaluation was clearly set out as a major device to pursuing quality assurance in education. The 2006 and 2008 guidelines provided schools with a basis for school evaluation and outlines for the implementation of self-evaluation and some kinds of external evaluation.

1.4. Achievements and problems associated with the implementation of school evaluation

School evaluation is mandatory in all schools in Japan, from kindergarten to high school. It is necessary to check that evaluations are done effectively, on both a political level and at a school level, thus fulfilling the purpose of the system. If school evaluation is not done effectively, it becomes a social waste; the situation must be improved.

One instrument used to measure the relative achievements of the school evaluation system was MEXT's

Table 3 The Three-Tier Model of School Evaluation

1) Self-evaluation
- carried out by all school staff under the leadership of the principal
- school goals, school plans and other plans are to be referred to

2) Evaluation by persons related to the school
- carried out by guardians, PTA and/or School Council members, other community members, persons related to articulated schools, etc.
- carried out based on the results of the self-evaluation

3) Third-party evaluation
- commissioned and delegated to an independent group of evaluators
- professional and objective evaluation from an external viewpoint

Source: Prepared by the author with reference to *School Evaluation Guidelines* of 2006.

"Survey on School Evaluation etc. Implementation". This survey was distributed to all schools, including national, public (prefectural and municipal), and private schools.[12] Surveys were conducted in 2006, 2008, 2011, and 2014. The clearest achievement during the initial years of school evaluation was its spread to almost all public schools. The 2008 survey following the amendment of the above-mentioned School Education Law, demonstrated that the percentage of schools that had conducted "self-evaluation" was as high as 92.4 percent. During the 2011 survey, total participation was as high as 96.7 percent, with 99.9 percent of public schools participating. School evaluation had become widespread nationally.

On the other hand, the survey shows the extent of the efficacy of practicing school evaluation. The surveys conducted in 2011 and 2014 asked the following questions:

"To what extent was the self-evaluation useful to both systematically and continuously improve educational activities or other school operations?"

The choices given were:

"a. extremely effective,
b. effective to some extent,
c. not very effective,
d. not effective at all,
e. don't know."

The answers for each year are shown in Table 4.

Table 4 Schools' responses to the usefulness of school evaluation

Q: "To what extent was the self-evaluation useful to both systematically and continuously improve educational activities or other school operations?"				
answers	2006	2008	2011	2014
a. extremely effective	-	-	16.3%	20.3%
b. effective to some extent	-	-	79.3%	74.1%
c. not very effective	-	-	2.1%	2.4%
d. not effective at all	-	-	2.3%	3.0%
e. don't know	-	-	0.1%	0.1%

Source: Author's translation of the MEXT's "Survey on School Evaluation etc. Implementation"
Letters have been added by the author.

For both of the years' surveys, the sum of the positive answers to a and b was around 95 percent. One should take that school evaluation was effective at improving schools' activities. However, we must understand that the survey was submitted by schools to their administrators or superiors. Therefore, regardless of the efficiency or the burden created by conducting the evaluation, it is very likely that principals would mark b instead of c if there were any tiny positive consequences of school evaluation . Therefore, we should be cautious about coming to optimistic conclusions from the percentages observed here.

However, other research studies could lead us to a more realistic view of the efficacy of school evaluation.

The material prepared by Nomura Research Institute (NRI), Ltd. (2011) included interesting results from their original questionnaire survey distributed to schools participating in the MEXT's school evaluation promotion conference in July and August 2009. They collected 428 responses. Their survey included the following item: "School evaluation leads to the improvement of school operations or activities and is reflected in students' academic performance and behavior, or other concrete effects."

The choices given and the percentages in the bracket were:

"a. extremely applicable (9.8),
b. applicable to some extent (65.9),
c. not very applicable (22.2),
d. not applicable at all (1.2),
e. no answer (0.9)."

An evaluation of the efficacy of school evaluation was measured by another survey item as follows:

"School evaluation leads to the reinforcement of staff cooperation and teamwork, or the improvement of their motivation."

The choices and the percentages in the bracket were:

"a. extremely applicable (7.2),
b. applicable to some extent (54.0),
c. not very applicable (35.3),
d. not applicable at all (2.1),
e. no answer (1.4)."

Following these percentages, the authors of the material explain, "School evaluation practices that lead to positive results score about 6 to 8 out of 10". They point out also the sample schools are "model schools" that study school evaluation under the authorization of MEXT, and ordinary schools may perform more poorly than "model schools" in making positive use of school evaluation.

In parallel with the study carried out by NRI, academics have been discussing the merits, achievements, and problems associated with school evaluation. In the record of the first meeting of the Working Group on What School Evaluation Should Be held on August 4, 2011, we see that this group had begun to make criticisms of school evaluation, such as: "Schools barely conduct evaluation but can't afford to make use of it for improvement" and "The reality is that the school evaluations conducted in many schools are not yet utilized."[13]

Takashi Ebisui was dispatched from his teaching profession in a public school to the National Graduate Institute for Policy Studies to conduct a study of school evaluation systems in different local educational authorities. His experience as a teacher led him to think about the reasons for the high percentage of positive answers to the 2011 survey despite his perception that begged the question "why is my actual feeling that school evaluation is not utilized enough for school improvement?" Thereafter, he noticed the material prepared by the Nomura Research Institute for the government working group at the request of MEXT dealt with his question; "Is it really advantageous to conduct school evaluation?" (Ebisui 2012)

Not that all cases are negative, as positive effects and good practices are reported, including that "teachers' morale was stimulated, and students and parents' feelings was understood." For instance, a lower-secondary school principal from Hiroshima city said, at the 2nd meeting on September 12, 2011 that "as evaluation items and the results were specific and clear, they could be utilized in school improvement." At the same meeting, an elementary school principal from Iwata City, Shizuoka Prefecture commented that "at first, effort indicators and performance indicators were confused, but we came to realize the importance of expressing goals as the children's accomplishment through faculty development and our practice." At the 3rd meeting of the Working Group on What School Evaluation Should Be on September 28, many good practices were introduced as "good examples" of the implementation of school evaluation.[14] However, these positive sides are not observed as widely as expected throughout Japan.

We shall return to the interpretation of Table 4. Following the results of the 2011 survey, MEXT commented that, "in sum, 95.6% of the schools answered positively in terms of the self-evaluation's usefulness in both systematically and continuously improving educational activities or other school operations." Yet, the MEXT added a reminder also noting that the "rate of schools choosing *extremely effective* is only 16.3%, and there is thus a task to enhance the effectiveness of school evaluation." Negative comments by MEXT itself appeared for the first time in these surveys. It is likely that the arguments at the panel or working group established within MEXT had some influence on the ministry's judgment.

Currently, many local educational authorities discuss openly the dysfunctional elements of school evaluations. Such matters are often discussed under themes such as "ensuring the effectiveness of school evaluation", especially after the said Working Group had used the phrase "promoting effective school evaluation" in the title of its report published in 2012. It became a national trend to pursue "effective school evaluation", that is, a general feeling that the school evaluation system was dysfunctional became commonly understood nationally.

Thus, this note focuses on the difficulties of school evaluation in Japan. From a solution-seeking position, the next section examines the issues that may obsturct the successful implementation of school evaluation.

2. Seven factors that may obstruct the successful implementation of school evaluation

2.1. The ambiguous meaning of the Japanese word for "evaluation"

In this section, the author has made some a number of reflections on the special features of Japanese school evaluation along with the key points of the system and the realities of implementation.

The first point is that the Japanese words "hyoka" (evaluation) and "hyoka-suru" (evaluate) are ambiguous. These words are sometimes used to mean different things. As the English verb "value" sometimes means that something is held to be of some worth, the Japanese word "hyoka-suru" often bears the meaning of "praise." Thus, the meaning of school evaluation may turn out to be about telling a school if it is bad or good. It would be hard to describe school evaluation as aimed at "improving school operation" or "improving the level of education" unless the evaluation is done in a constructive manner.[15]

2.2. School evaluation not linked to the school's goals

The Guidelines, as stated above, proposed that schools should refer to their annual goals, as set for each academic year, in carrying out their self-evaluation. Schools should evaluate if their goals were met, if their strategies and practices to meeting their goals were adequate, and note if there was anything to improve.

The report of the Working Group on What School Evaluation Should Be (2012) was issued under the title of *Creating schools with the local community and promoting effective school evaluation (Report)*. The report recommended that, for effective school evaluation, schools, "clarify and emphasize goals." After the national government announced the need for systematic improvement in the efficacy of school evaluation, local governments pursued improvements in their school evaluation policy.

For instance, the Oita Prefectural Board of Education deepened its school evaluation policy to utilize school evaluation in school operation. Its Schools' Collaborative Problem Solving Ability Improvement Study Committee published a report titled *On Improvements to Schools' Collaborative Problem Solving Ability (Proposal)*[16] on September 20, 2012. In the report, the state of the relationship between school evaluation and school goals was described as follows (letters have been added by the author):

a. The school's goals lack specificity, so it is hard for school staff to come to a common understanding and to accomplish projects requiring teamwork.
b. Specific school goals, as numerical goals, are not fixed, and, therefore, school evaluation would not effectively contribute to school improvement.
c. Specific school goals and (evaluation) results and improvements are not made public, so establishing collaboration with parents and the community members is difficult.

After pointing out shortcomings in the goal-setting process, the Oita Prefectural Board of Education redesigned the school evaluation process to be simpler, and much closer to the school planning process, making school evaluation more effective (Oita Prefectural Board of Education 2013).[17]

However, not all local authorities acted in the same way as Oita, though they may have experienced similar problems. Even schools that were invited to speak at the above-mentioned Working Group were far from perfect. One principal of a municipal lower-secondary school in Hiroshima City, an advanced school in practicing school evaluation, reported the school's practice and was then asked a question by a Working Group member: "While you talked about schools with their school goals not being specific, your school's short-term goals are also the same except some of the prioritized goals. What do you say?" The principal then confessed the difficulty of establishing goals through the broad participation of teachers noting, "I think it is also a good idea to leave setting short-term goals in teachers' hands, but we can't afford to take that time, and so I make the plan myself."[18]

In addition to these documents, according to schools' websites, many school goals in public elementary and lower high schools are not specific in terms of targets for student achievement. For instance, in Chuo Ward, 38 out of 52 Schools' goals were difficult to evaluate, as they were not focused on student achievement (Hashimoto 2017).

Unless goals are "S.M.A.R.T." (Specific, Measurable, Achievable, Realistic, and Time-limited), evaluation cannot be effective. School evaluation can be done effectively when school goals are designed with this intention in mind. Practices such as the Oita Prefectural Board of Education's "cooperative goal achievement" movement by the schools, family, and community, should be studied to set school goals that will penetrate school plans, school practices, school evaluation, and school improvement.[19]

2.3. Poor indicators for evaluation

This matter is clearly observed in the results of MEXT's "Survey on School Evaluation etc. Implementation" from 2008 and 2011.[20] In the 2008 survey, it was reported that 46.9 percent of the prefectures and 29.5 percent of the municipalities (wards, cities, towns, and villages) had established common "evaluation items and/or indicators for school evaluation." Most compulsory education schools are under the jurisdiction of municipalities; thus, over 70

percent of schools must consider setting their own indicators.

Principals appear to be struggling with setting convincing indicators for measuring the performance of their activities. In the 2011 survey, there was a question that asked respondents to "choose the matters, if any, that make you think of a problem or difficulty concerning self-evaluation (multiple responses permitted)." Out of the eleven options, the top three chosen were (with the respondents' selection rate in brackets):

1. Setting evaluation items and evaluation indicators (38.9%)
2. Feelings of fatigue among teachers and staff (36.8%)
3. Utilization of evaluation results (26.4%)

However, we must be careful in making conclusions from these data. We may not conclude that the rest of the schools have similar understandings of their evaluation indicators. In the same survey, we can see the types of indicators actually used. The answers to a question that asked schools what kind of indicators they used are as follows (letters added by the author):

a. Results of a questionnaire survey for parents: 78.5%
b. Results a questionnaire survey for students: 62.0%
c. Seeking comment from persons related to the school: 54.5%
d. Matters related to student guidance: 53.6%
e. Results of academic ability surveys: 53.5%
f. Results of physical ability and strength surveys: 42.2%
g. Seeking comment from parents: 38.4%
h. Results of the questionnaire survey for persons related to the school: 29.1%
i. Conditions of career education and guidance: 24.3%
j. Conditions of extra-curricular activities: 23.6%
k. Seeking comment from students: 18.9%

The top two indicators hold an absolute majority. Even after hearing about "good" examples, a Working Group member pointed out that these examples relied too much on results of the questionnaire surveys posed to students in checking their "academic achievement."[21] Nationwide, school websites show that actual indicators are mostly a list of what teachers should do.[22] Schools are failing to set indicators that reflect students' performance and other indicators that will stand for the school's real achievement.

2.4. Evaluation results not utilized for school improvement

This matter is also clearly observed through the survey results. In the above-mentioned question on "the problem or difficulty concerning self-evaluation (multiple responses permitted)" in the 2011 survey, the third most chosen option was the "utilization of evaluation results," at 26.4 percent.[23] It is then clear that more than a quarter of schools are concerned about whether they are good at utilizing the results of evaluation. Again, however, we must be careful about the interpretation of the percentages, so that we do not conclude that the remaining schools are good at or are confident in the usage of the evaluation. In the same survey, responses to a question that asked about "ways of utilizing the results of self-evaluation" were as follows (letters added by the author):

a. Share evaluation results through explanations at staff meetings, etc. (90.0 %)
b. Set an opportunity to discuss strategies for improvement, following the evaluation results, at staff meetings, etc. (78.2 %)
c. Share evaluation results through explanation to parents and community members, etc. (36.1 %)
d. Set an opportunity to discuss strategies for improvement following the evaluation results to parents and community members, etc. (17.6 %)
e. Other (4.1%)

Most schools appear to regard "explanation" and "discussion" as typical ways of utilizing evaluation results. However, "explanation" and "discussion" are mere passages to true utility. These responses suggest that schools understand utilization at the effort level, and there could be a fear that the aim of school evaluation will end at an intermediate level before desirable effects are achieved.

The summary of the proceedings of the Working Group on What School Evaluation Should Be offers some "good examples," including the case of Musashimurayama City in the western suburbs of Tokyo. It is reported that the Board of Education and the municipal schools would check the results of education practices along the given indicators, and, following the "intermediate evaluation which is an self-evaluation," they would begin to plan the next fiscal year's school budget. At the end of the fiscal year, schools would endeavor to disclose evaluation results by "clarifying what was fruitful and what was problematic and proposing ways to improve the problematic matters." The Board of Education would "make sure that concrete improvement measures following school evaluation are clearly outlined when setting each school's curriculum for the next school year."[24]

The practices of Musashimurayama City have not been meta-evaluated so I will not judge whether they are really "good" practices or not; however, I will note that municipalities that intentionally link school evaluation and the support of the board of education toward with school improvement are not at all popular. School evaluation results that are related to the next year's curriculum planning could scarcely be seen in other cities, even if there is no research on this question. Even after listening to the case of Musashimurayama City, a member of the Working Group noted that "all schools' evaluation results are brought to the boards of education, but the boards are not sure how to process these results. It is difficult to determine how the results should be reflected in the making of budget plans for the next year, which is almost left untouched.[25] Thus, it must be understood that school evaluation is not utilized widely enough in many schools.

2.5. An absence of appropriate evaluators or coordinators

No full-time staffs are assigned for school evaluation in each school and, barely any in boards of education. Even if a person has been specifically hired for school evaluation, they are likely to be non-regular staff and non-specialists in evaluation. As for external evaluation, such as the "evaluation by persons related to the school," this has been left to the discretion of schools while self-evaluation is done by the principal, the teachers, and others.[26] According to the 2011 survey, schools are appointing evaluators as follows (letters added by the author):[27]

a. PTA headquarters member (67.5%)
b. School council member (61.2%)
c. Residents' association member (40.0%)
d. Social welfare facility or organization member (31.9%)

e. Former school council member (31.3%)
f. Social education facility or organization member (19.7%)
g. Parents other than PTA headquarters members, etc. (19.3%)
h. People of experience or academics (13.9%)
i. Former school staff (10.4%)
j. Staff from other schools (9.6%)
k. Alumni (8.9%)
l. Concerned persons from local firms or NPOs (5.9%)
m. Coordinator or volunteer from a school support organization (5.5%)
n. School council member (4.5%)
o. Councilor from a private education institution (4.3%)
p. Others (3.8%)

Most of these people may be in positions to understand the school. However, they are not, in most cases, well trained in evaluation or prepared to evaluate schools. In the Working Group's arguments, there were some comments that "even though it depends on the member, it is a problem, in reality, that proper evaluation is usually difficult to conduct."[28] Again, properly trained evaluators or coordinators for school evaluation are rarely provided.

2.6. Lack of training courses

Training courses concerning school evaluation are insignificant in both the pre- and in-service training received by teachers. During pre-service, mainly carried out in teacher training universities and colleges, there is sometimes a class on school evaluation. One example may be a half-year (one semester) lecture on "Educational Administration and Finance" for third year students at the Faculty of Education at Hirosaki University, a typical local national university in northern Japan. Out of its 15 classes, the lecture on "School Administration" is in the 12th class, and that on "The relationship between the school and parents, students and community members" is in the 13th class.[29] In the same university's advanced program, there is another half-year (one semester) seminar on "Issues and practices of education management" for first year Masters students. Out of its 15 classes, the "School evaluation and school improvement" lecture takes place in the 14th class. These seem to be the only chances for the students to examine school evaluation while in university.[30]

In general, there could be at most only one timeslot allocated on a single day of a series of lectures titled "educational administration" or "school management." In many cases, there may not be appropriate instructors to lecture on school evaluation, and in those cases, students do not gain a realistic sense of the situation but merely learn about school evaluation through books.

In-service training about school evaluation is provided mainly in prefectural teacher training institutes or, in some large cities, municipal teacher training institutes. Lectures and seminars on school evaluation are provided mainly as partial supplements in one course or program as illustrated in Table 5.[31] In some cases, the school evaluation curriculum is well contextualized with related material such as school goals and school management.

2.7. Poor funding for evaluation

The problems raised above arise from, or result in, the lack of an appropriate budget for school evaluation. Schools and local authorities are not able to hire evaluators, coordinators, mentors, or assistants to lead or assist in school evaluation. At the very early stages of designing the school evaluation system, there was some hope that budgets would be allocated for school evaluation as it became mandated. The abovementioned Special Committee for

Table 5 Samples of teacher training programs with school evaluation curriculum (2012-13)

Hosted by	Title of the program	Time allocated for school evaluation	Remarks
Tokyo Metropolitan Board of Education	School supervisor initial training program	2 hours	
	Public school principal candidate training program	2 hours	
	School evaluation leader training program	3 days	2012, in collaboration with JES
Hiroshima City Board of Education	School evaluation system evaluator training program	2 days	2013, in collaboration with JES

Source: Prepared by the author following Hashimoto 2014

Compulsory Education, set up under the Central Council for Education, noted, "it is proper to make it [self-evaluation] compulsory. Because, if it becomes compulsory, allowance will be provided for it, won't it."[32]

However, costs for evaluation were scarcely considered in the school policy-making process. For instance, in the abovementioned Panel of Research Experts for Promoting School Evaluation appointed in July 2006, information on evaluation costs was presented by outside informants. A person in charge of higher education at MEXT responded to a Panel member's question about the costs of university evaluation by noting, incorrectly, "some million yen is to be paid to the evaluation agency" and that money would include "fees for evaluators, costs to hold evaluation meetings and travel costs and so on for site visits." Likewise, a childcare researcher stated, at the same meeting, "for the childcare center's third-party evaluation, the budget varies from 400 to 700 thousand yen per center, and that is a quite tight budget for the evaluation agency which must at least secure three evaluators per center."[33]

However, the members did not devote much time to discussing the budget problem with the exception of some members, who mentioned that "the third-party evaluation system in Shinagawa Ward requires a million and some hundred thousand yen per year, if it is to evaluate all of its 58 elementary or lower-secondary schools within five years. It may be financially difficult to design a nationwide system of third-party evaluation in Japan. Nevertheless, if we must do it, we need to investigate conditions that will enable it."[34] Another member stated, "conducting assessments of academic ability and third-party evaluation training etc. should be costly. We must also think about the cost problem."[35] These comments are reported on the Panel's website.

In sum, there are few budgetary comments beyond those about third-party evaluation, and even in this case, a view like a member's comment that "in addition to the possibility of realization, the largest problem is the question of the necessity for the nation to do it in all schools"[36] was the representative comment. The national councils, panels and working groups discussed in this paper did not stress the importance of setting aside a portion of budgets for the evaluation of schools. Thus, budget was not allocated with the approval of the school evaluation system.

The lack of funding was not only a problem within the evaluation process but also within the post-evaluation process; that is, little funding was saved for improvement following the results of the evaluation. At the early stages of the implementation of school evaluation, as observed in the proceedings of the 2006-2007 meetings of the Panel of Research Experts for Promoting School Evaluation, there are significant debates about necessary follow-up treatment carried out by local authorities for schools following the evaluation results. Some years later, however, the situation is not optimistic, as the comments cited above (2.4) remain critical. It is argued that the design of the school evaluation system was made insufficient in terms of transparency concerning schools' planning and practice processes, which may have undermined the basis for supporting schools financially following the evaluation

results.[37] School evaluation experts and policy-makers in Japan have a great deal to consider going forward, including system design and its relationship to budgeting.

3. Conclusion: Reflections on the special features of the Japanese system of school evaluation

This note consists of an outline of Japanese school evaluation system and discusses the assumption that it is insufficiently effective and requires additional components at the implementation stage. The outline itself is not an original finding, and the supposition is as yet unproven. The presentation and combination of elements, however, contribute to the originality of this note. In conclusion, I shall suggest the direction of further study from an international perspective.

3.1. The circumstances surrounding Japanese school evaluation

The seven topics I examined in this paper were chosen rather subjectively with each subject being investigated in very few research studies. A need for further research in this area remains, although their problematic context has been reviewed, and many important viewpoints have been elaborated by education experts in the national councils and panels mentioned in this note. School evaluation in Japan lacks these seven elements, but most lacking may be the context and connections between these seven elements. There is also a strong need and great opportunity to investigate these context and connections further.

Beyond these seven concrete issues, the two contexts of school evaluation policy and the wider education system ought to be taken into consideration. One area is the domestic context of educational policy. As the promotion of school evaluation began under political pressure, and was conducted under the promotion of education administrations, the basic strategies of school evaluation changed occasionally, and were strongly affected by the political trends and public opinions of the time. The current trend is around "creating schools with the local community." As an aforementioned report's title suggests, promoting "effective school evaluation" tends to be linked to the subject of "creating schools with the local community" This subject is of course also important, but, as researchers, we have to reflect the purpose of school evaluation and the connections between the factors discussed in this paper.

The other context involves international trends. Those countries that have a school evaluation policy have shifted politically, from a strong emphasis on external monitoring and control, toward greater school and teacher autonomy, consisting of capacity building and self-evaluation.[38] Based on some European case studies, McNamara and O'Hara explain the reasons for this trend: first, that "context and tradition are still influential in educational policy" and "teaching and teachers are still highly regarded" in many European countries. Secondly, despite the vigorous external evaluation developed in England and the United States, there is "some evidence that the limitations and side effects (particularly in relation to teacher morale and retention) is resulting in a rethink." Thirdly, by the national government's range of school evaluation objectives, "additional complexity and ambiguity is added to school evaluation". In addition to these three, the problem of "the high cost of external inspection" has also been pointed out (McNamara and O'Hara 2009:275).

Japanese school evaluation system is developing historically still, and is possibly on the edge of remaining a mere facade. However, from a political viewpoint, as in many other countries Japanese school evaluation policy was initiated in the context of New Public Management strategies, which entailed a shift in focus from external review processes to self-evaluation. The latter point shall be examined in the next sub-section.

3.2. Toward an understanding of Japanese school evaluation in an international context

Finally, I shall suggest directions for further research in an international context, where self-evaluation presides over external evaluation in school evaluation policies.

This semi-special issue of "School evaluation studies from international perspectives" was originally planned to compare school evaluation policies and practices in Korea, Taiwan, and Japan, with those of Victoria Province in Australia. The "Whole School Evaluation (WSE)" process was the focus of interest, as discussed in John Owen's article[39]. We could not accomplish a comparison across all the four systems, as we lacked the necessary criteria for comparing; the common understanding of the main characteristics of WSE. At the end of this paper, however, I hope to establish the importance of the comparison by considering whether Japanese school evaluation follows WSE processes.

In this semi-special issue, John Owen's article refers to Sanders, J.R. & Davidson, E.J. (2003), and say "WSE can be defined as the use of systematic investigation of the quality of a school and how well it serves the needs of its community". We can find similar but different expressions of the definition of WSE in articles that examine cases in Ireland and South Africa.

A report of OECD, written by using Ireland Department of Education and Science (DES) official website information, describes WSE as "a 'whole school' focus in school evaluation" (OECD 2007:61). In detail, it also says "Whole School Evaluation is a process whereby a team of Inspectors from the DES spends a few days in a school evaluating the overall work of the school under the following themes", enumerating qualities of "school management", "school planning", "curriculum provision", "learning and teaching in subjects", and "support for students" (OECD 2007:13).

A report by the OECD, written using Ireland's Department of Education and Science (DES) official website information, describes WSE as "a 'whole school' focus in school evaluation" (OECD 2007:61). In detail, it says also, "Whole School Evaluation is a process whereby a team of Inspectors from the DES spends a few days in a school evaluating the overall work of the school under the following themes", enumerating qualities of "school management", "school planning", "curriculum provision", "learning and teaching in subjects", and "support for students" (OECD 2007:13).

Papers written in South Africa make many mentions also of WSE. During his conference paper of 2001,[40] Mgijima mentioned, "Whole School Evaluation is one intervention to move schools that are in a critical situation along the path of becoming effective." Later, Madikida (2016) refers to Govender, Grobler & Mestry (2015), by defining WSE as "a quality assurance system aiming at improving quality teaching and learning in schools. It focuses on improving the overall quality of education in South Africa by means of internal and external evaluations".

When we assess Japanese school evaluation in terms of these definitions of WSE, we may say it corresponds to WSE. Japanese school evaluation involves a "whole school" focus in school evaluation, and it uses a similar range of evaluation items as the WSE process conducted in the countries mentioned above.

On the other hand, Japanese school "self-evaluation" seems to stand alone with little connection made between school goals, plans, and actual strategies to enrich the educational process. As mentioned at the beginning of the paper, the Japanese school evaluation system necessitates that "school goals, school plans and other plans are to be referred to";[41] however, is likely to undervalue those goals and plans in actual practice. Instead, the process seems more like the "use of systematic investigation of the quality of a school", as defined in Owen's article. What constitutes "systematic", however, remains a critical problem, and the investigation of "how well it serves the needs of its community" is open to question. In conclusion, we must not simply regard Japanese school evaluation as one type of WSE, as it does not adhere to the WSE framework strictly enough.

This note will propose nothing further. I will only point out that studying WSE in other countries and regions will enrich research investigating Japanese school evaluation, and possibly Korean and Taiwanese school evaluation studies if they have common elements warranting further study. International comparisons will assist us in noticing the many variations and patterns between those connections and between different factors. It is regrettable that I cannot extend my research to such comparative investigations, but JES and other societies and associations for the study of evaluation should be open to these tasks.

Acknowledgments

The author would like to thank the editors for the review with constructive comments. Helpful English language review was provided by Enago.

Notes

1 Ministry of Education, Science and Culture became MEXT under the reorganization of the ministries and agencies of 2001.

2 MEXT (Ministry of Education, Culture, Sports, Science and Technology), *Gakkō kyōiku hō sekō kisoku* [Enforcement Regulations of the School Education Law], Ministerial Ordinance No. 14 of 2002. It came into effect in April 2002.

3 The Council for Regulatory Reform was established under the jurisdiction of the Cabinet Office, active from April 1st 2001 to March 31st 2004.

4 *Gakkō kyōiku hō* [School Education Law], Law No. 26 of 1947, amended by Law No. 96 of 2007, art. 42. Even though it mentions only 'elementary school', the article shall be applied mutatis mutandis to other schools including kindergartens.

5 MEXT, *Gakkō kyōiku hō sekō kisoku* [Enforcement Regulations of the School Education Law], Ministerial Ordinance No. 34 of 2007.

6 Minutes of the 13th meeting of the Council on Economic and Fiscal Policy.

7 *On the School Evaluation System Study Group* on the MEXT website.

8 The summary minutes of the 1st meeting of the School Evaluation System Study Group held on February 20th 2006, and that of the 2nd meeting.

9 See Hirota and Ikeda 2009 for detail.

10 MEXT. 2007. *On the Principle and Future Measures for Promotion of School Evaluation; The First Report* was released in August 2007. The main points are put on the website.

11 This word was already publicized in MEXT Ministerial Ordinance No. 34 of 2007. Refer to note 5.

12 Survey on School Evaluation etc. Implementation is conducted every three years. The survey form was distributed to all schools, including national, public (municipal), and private, headed to each school principal. The top page of the website for the survey's URL is http://www.mext.go.jp/a_menu/shotou/gakko-hyoka/1322262.htm (Retrieved December 11, 2017).

13 The Working Group on What School Evaluation Should Be was established in August 2011 within the Panel of Research Experts on the Desirable Improvements for School Operation, The summary minutes of the 1st meeting of the Working Group can be viewed at the following URL;
http://www.mext.go.jp/b_menu/shingi/chousa/shotou/078_1/gijigaiyou/1309993.htm (Retrieved December 11, 2017).

14 The summary minutes of the 2nd and 3rd meetings of The Working Group on What School Evaluation Should Be can be viewed at the following URL; http://www.mext.go.jp/b_menu/shingi/chousa/shotou/078_1/gijigaiyou/1311974.htm (Retrieved December 11, 2017).
http://www.mext.go.jp/b_menu/shingi/chousa/shotou/078_1/gijigaiyou/1312237.htm (Retrieved December 11, 2017).

15 There seems to be no linguistics research for this item; my personal experiences are reflected in this statement.

16 Schools' Collaborative Problem Solving Ability Improvement Study Committee (Oita Prefecture). 2012. *On the Schools'*

Collaborative Problem Solving Ability Improvement (Proposal).

17 This guidance for school evaluation is retrieved at;
http://www.pref.oita.jp/uploaded/attachment/2003491.pdf (Retrieved December 11, 2017)

18 The URL is as same as the note 14.

19 Information can be earned at Oita Prefecture's website; http://kyouiku.oita-ed.jp/kikaku/2016/05/28-2.html (Retrieved December 11, 2017).

20 Same as note 12, *supra*.

21 The summary minutes of the 4th meeting (held on October 5th 2011) of The Working Group on What School Evaluation Should Be.

22 See Hashimoto 2017 that points out this is actually seen in school's websites in Tokyo.

23 Same as note 12, *supra*.

24 Same as note 21, *supra*.

25 *ibid.*

26 Same as note 5, *supra*.

27 Same as note 12, *supra*.

28 The summary minutes of the 2nd meeting of The Working Group on What School Evaluation Should Be (held on September 12th 2011). The URL is as same as in note 14.

29 Hirosaki University, The Faculty of Education [2017], Syllabus.

30 Hirosaki University, The Graduate School of Education, Educational Practice Program [2017], 2017 Syllabus.

31 See Hashimoto 2014. In addition to prefectural and municipal training institutes, the National Institute for School Teachers and Staff Development offers lectures and exercises (http://www.nits.go.jp/en/ L.A. December 11, 2017). The Japan Evaluation Society also conducts lectures and exercises about school evaluation that are not only geared toward teachers (http://evaluationjp.org/english/index.html L.A. December 11, 2017).

32 The minutes of the 33rd and 34th meeting of the Special Committee for Compulsory Education held on September 8th 2005. Minutes of this Committee could be retrieved at the MEXT's website (http://www.mext.go.jp/b_menu/shingi/chukyo/chukyo6/index.htm).

33 The summary minutes of the 4th meeting of the Panel of Research Experts for Promoting School Evaluation held on October 30th 2006. Minutes of this Panel could be retrieved at the MEXT's website (http://www.mext.go.jp/b_menu/shingi/chousa/shotou/037/old_index.htm).

34 The summary minutes of the 8th meeting of the Panel of Research Experts for Promoting School Evaluation held on February 27th 2007. The URL is as same as the note 33.

35 The summary minutes of the 15th meeting of the Panel of Research Experts for Promoting School Evaluation held on October 30th 2007. The URL is as same as the note 33.

36 The summary minutes of the 13th meeting of the Panel of Research Experts for Promoting School Evaluation held on August 27th 2007. The URL is as same as the note 33.

37 Opinion seen in Miura 2010.

38 This paragraph is written referring to McNamara and O'Hara 2009. See their work for individual articles they had referred to.

39 Refer to John Owen's paper of " Whole School Evaluation: Approaches Used in the Public School System in Victoria, Australia" in this volume.

40 In Risimati (2007:6), we can see quotes from Mgijima's conference paper of "The South African hmodel for Whole School Evaluation" issued in 2001.

41 *School Evaluation Guidelines*, 2006

References

Cabinet Office. (2003). *The Third Report of the Council for Regulatory Reform* (in Japanese). Tokyo: Cabinet Office.

Cabinet Office (2005). *Basic Policies for Economic and Fiscal Management and Structural Reform 2005* (in Japanese). Tokyo: Cabinet Office.

Central Council on Education. (1998). *Report on How the Local Administration on Education Should Be* (in Japanese). Tokyo: Ministry of Education, Science and Culture.

———. (2005a). *Creating a New Era of Compulsory Education* (in Japanese). Tokyo: Ministry of Education, Culture, Sports, Science and Technology.

———. (2005b). *Structural Reform of Compulsory Education* (in Japanese). Tokyo: Ministry of Education, Culture, Sports, Science and Technology.

Ebisui, Takashi. (2012). "Inter-municipal Comparative Study of Administrative Management for School Evaluation System Improvement: Comparative Analysis of Kami, Shinagawa, and Yokohama Cities" (In Japanese). A policy paper for the master's degree at National Graduate Institute for Policy Studies. Retrieved December 27, 2017 (http://www3.grips.ac.jp/~education/epc/wp-content/uploads/2014/04/201103.pdf).

Furuta, Kaori. (2013). "A Study of the Policy Making Process of Independent Evaluations of Schools" (In Japanese). *Bulletin of Kyoto Seibo College* 42:53-71. Retrieved December 27, 2017 (https://ci.nii.ac.jp/naid/110009577990/).

Govender, Neelam, Bennie Grobler, and Raj Mestry. (2015). "Internal Whole School Evaluation in South Africa: The Influence of Holistic Staff Capacity." *Education Management, Administration and Leadership Journal* 44(6):1-25.

Hashimoto, Akihiko. (2009). "Accountability Systems in Japanese Education; a Note on the National Testing and the School Evaluation Policies." Pp.59-73 in *Global and Comparative Perspectives in Academic Competence, Evaluation, and Quality Assurance, The 10th International Conference on Education Research (ICER)*. Seoul, Korea: Seoul National University.

———. (2011). "The Design, Implementation, and Operation of an Evaluation System that Works in Modern Educational Supervision." Pp.37-48 in *Bulletin for the Beijing International Forum on Education Supervision and Evaluation 2011*. Beijing China: Beijing City Office.

———. (2012). "The Establishment of the 'Certified Specialty Evaluators Training Program for School Evaluators' and its Perspective" (in Japanese). *Japanese Journal of Evaluation Studies* 12(1):15-25. Retrieved December 27, 2017 (https://www.jstage.jst.go.jp/article/jjoes/12/1/12_1_15/_article).

———. (2014). "School Evaluator Development Programs: Implementation of Evaluation Training at the Local Education Authority" (in Japanese). Pp.33-38 in *Proceedings: The 11th Spring Conference of the Japan Evaluation Society.*

———. (2017). "Patterns of Implementing School Evaluation - Investigating Websites of Public Schools in Tokyo" (in Japanese). Pp.125-128 in *Proceedings: The 14th Spring Conference of the Japan Evaluation Society*.

Hirota, Teruyki and Masanori Ikeda. (2009). "Political Process over Institutionalization of School Evaluation in Japan: from the Latter Half of the 1990's to 2006" (in Japanese). Nihon University, The Institute of Humanities and Social Sciences *Kenkyu Kiyo* 77: 39-66.

Katsuno, Masaaki. (2016). *Teacher Evaluation Policies and Practices in Japan: How Performativity Works in Schools*. Routledge Series on Schools and Schooling in Asia. Oxford, UK: Routledge.

Katsuno, Masaaki and Tetsuro Takei. (2008). "School Evaluation at Japanese Schools: Policy Intentions and Practical Appropriation." *London Review of Education* 6(2):171–181. Retrieved December 27, 2017 (https://www.researchgate.net/publication/240517418_School_evaluation_at_Japanese_schools_Policy_intentions_and_practical_appropriation).

Kioka, Kazuaki. (2005). "A Review on Theory and Practice of the School Evaluation in Japan - Focusing on the Conflict of Organizational Control and Organizational Development" (in Japanese). *NIER Research Bulletin* 134:59-80. Retrieved December 27, 2017 (https://www.nier.go.jp/kankou_kiyou/kiyou134-all.pdf).

Koyama, Nagaki. (2008). "Educational Administration in Japan and the Role of Local Governments." Papers on the Local Governance System and its Implementation in Selected Fields in Japan No.9. Tokyo: Council of Local Authorities for International Relations and National Graduate Institute for Policy Studies. Retrieved December 27, 2017 (www3.grips.ac.jp/~coslog/activity/01/04/file/Bunyabetsu-9_en.pdf).

Madikida, Pakama Patience Patricia. (2016). "Challenges in the Implementation of Whole School Evaluation at Secondary Schools in the Libode District, Eastern Cape Province." Dissertation for D. Ed. at the University of South Africa. Retrieved December 27, 2017 (http://uir.unisa.ac.za/handle/10500/21902).

McNamara, Gerry and Joe O'Hara. (2009). "Where Global Meets Local: Contexts, Constraints, and Consensus in School Evaluation in Ireland." Pp.273-290 in *The Sage International Handbook of Educational Evaluation*, edited by Katherine Ryan and J. Bradley Cousins. Thousand Oaks, CA, USA: Sage Publications Inc.

Ministry of Education, Culture, Sports, Science and Technology. (2006). *Guidelines for School Evaluation at Compulsory Education Stage* (in Japanese). Tokyo: Ministry of Education, Culture, Sports, Science and Technology.

———. (2008). *Guidelines for School Evaluation*. (in Japanese). Tokyo: Ministry of Education, Culture, Sports, Science and Technology.

———. (2010). *Guidelines for School Evaluation*. (in Japanese). Tokyo: Ministry of Education, Culture, Sports, Science and Technology.

———. (2016). *Guidelines for School Evaluation*. (in Japanese). Tokyo: Ministry of Education, Culture, Sports, Science and Technology.

Miura, Satoko. (2010). "Reflections on the System of School Evaluation for the Purpose of School Improvement" (in Japanese). *NIER Research Bulletin* 139:113-127. Retrieved December 27, 2017 (https://www.nier.go.jp/kankou_kiyou/kiyou139-013.pdf).

Nagao, Masafumi. (2007). "Theoretical and Practical Issues in School Evaluation" (in Japanese). *Japanese Journal of Evaluation Studies* 7(1):3-19. Retrieved December 27, 2017 (www.idcj.or.jp/JES/jjes7_1nagao.pdf).

Nomura Research Institute, Ltd. (2011). "Creating a School with Organizational Strength using School Evaluation" (in Japanese). A material for the 3rd meeting of the Working Group on What School Evaluation Should Be on September 28. Retrieved December 27, 2017 (http://www.mext.go.jp/b_menu/shingi/chousa/shotou/078_1/shiryo/__icsFiles/afieldfile/2011/10/05/1311629_2.pdf).

OECD Improving School Leadership Activity, Education and Training Policy Division. (2007). *Improving School Leadership: OECD Project Background Report – Ireland*.

Oita Prefectural Board of Education. (2013). "Guidance for School Evaluation" *Toward a 'Straight thinking organization' that will collaborate to achieve the goal* (in Japanese). Retrieved December 11, 2017 (http://www.pref.oita.jp/uploaded/attachment/2003491.pdf).

Oouchi, Susumu. (2013). "Practical Tasks and Approaches to School Evaluation Based on the Characteristics of Special Needs Schools." *Bulletin of The National Institute of Special Needs Education* 12:36-59. Retrieved December 27, 2017 (https://www.nise.go.jp/cms/resources/content/8770/20131028-190018.pdf).

Panel of Research Experts for Promoting School Evaluation. (2007). *On the Principle and Future Measures for Promotion of School Evaluation; The First Report* (in Japanese). Tokyo: Ministry of Education, Culture, Sports, Science and Technology. Retrieved December 27, 2017 (http://www.mext.go.jp/b_menu/shingi/chousa/shotou/037/toushin/07061901.pdf).

Risimati, Hasani Pius. (2007). "Whole School Evaluation in Rural Primary Schools in Limpopo Province." Dissertation for D. Ed. at the University of South Africa. Retrieved December 27, 2017 (uir.unisa.ac.za/handle/10500/2224).

Sanders, James R., & Davidson E. Jane. (2003). "A Model for School Evaluation." Pp.807-826 in *International Handbook of Educational Evaluation*. Boston, USA: Kluwer.

Schools' Collaborative Problem Solving Ability Improvement Study Committee [Oita Prefecture]. (2012). *On the Schools'*

Collaborative Problem Solving Ability Improvement (Proposal) (in Japanese). Oita: Oita Prefecture. Retrieved December 27, 2017 (http://www.pref.oita.jp/site/kyoiku/2000759.html).

Working Group on What School Evaluation Should be. (2012). *Creating schools with the local community and promoting effective school evaluation (Report)* (in Japanese). Tokyo: Ministry of Education, Culture, Sports, Science and Technology. Retrieved December 27, 2017 (http://www.mext.go.jp/a_menu/shotou/gakko-hyoka/05111601/1318815.htm).

(Accepted 4th January, 2018)

Publication Policy of the Japanese Journal of Evaluation Studies

Last revised on 15th February 2005

The Purpose and the Name

1. The Japan Evaluation Society (hereinafter referred to as "evaluation society") publishes "The Japan Journal of Evaluation Studies (hereinafter referred to as "evaluation study") in order to widely releaseevaluation studies and outputs of practical activities to domestic and international academic societies,interested individual and institutions, and contribute to the advancement and prevalence of evaluationpractice.

Editorial Board

2. The editorial board administrates editing of evaluation study based on the editorial policy stated below.
3. The editorial board is formed with less than 20 members of the evaluation society who are assigned by the board of directors. Terms of editors are two years but can be extended.
4. The editorial board assigns one editor-in-chief, two vice-editors-in-chief, and a certain number of standing editors among the members.
5. The editorial board may hold at least one meeting to discuss the editing policy, plans of editorial board, and others.
6. The editorial board reports activities to the board of directors as needed and receives approval. Also it is required to report the progress of the past year and an activity plan for the following year at the annual conference.
7. The editor-in-chief, the vice-editors-in-chief and the standing editors organize the standing committee and administrate editing on a regular basis.

Editorial Policy

8. The evaluation study, as a principle, is published twice a year.
9. The evaluation study is printed on B5 paper, and either in Japanese or English.
10. Papers published in the evaluation study are categorized as five types;

10.1. Review
10.2. Article
10.3. Research note
10.4. Report
10.5. Others

11. The qualified contributors are members of the evaluation society (hereinafter referred to as "members") and persons whose contribution is requested by the standing editors. Joint submission of members and joint submission of non-members with a member as the first author are accepted. Submission by the editors is accepted.
12. Submitted manuscripts are treated as the above categories, however, the standing editors will decide based on the application of the contributors and the following guidelines;

12.1. "Review" is a paper, which provides an overview of evaluation theory or practice. The editorial board will make the decision regarding publication.
12.2. "Article" is considered as a significant academic contribution to the theoretical development of evaluation or understanding of evaluation practice. The standing editors committee makes adoption judgments following the referee-reading process described in the next section.
12.3. "Research note" is a discussion equivalent to the intermediate outputs of a theoretical or empirical enquiry. The standing editors committee makes adoption judgments following the referee-reading process described in the next section.

12.4. "Report" is the study report related to a practical evaluation project or evaluation. The standing editors committee makes adoption judgments following the referee-reading process described in the next section.

12.5. "Others" includes requested papers for special editions organized by the editorial board and announcements from editorial board to members regarding publication.

13. The editorial board selects two referee readers. For the "article" , the editorial board makes adoption judgments referring to the results from referee readings and comments provided by one editor assigned by the editorial board. For "review", "research note", "report" and "others", the editorial board makes adoption judgments referring to the results from referee readings.

14. When editors submit a manuscript, the editors are not allowed to attend any of the standing editors committee meetings or editorial board meetings regarding the manuscript.

15. The standing editors have alternative of approval or not-approval for adoption judgment of manuscripts submitted to any categories. However exception is permitted if the editorial board approves the publication after minor rewrite. Even if the manuscripts are considered insufficient as an "article", standing editors can decide whether the manuscripts are published as a "research note" or "report" if the authors wish to publish.

Formulation and Release of Submission Procedure

16. The editorial board formulates the submission procedure based on the editorial policy described above and release after approval from the board of directors.

Distribution

17. The evaluation study is distributed to all members for free and distributed to non-members for a charge.

Distribution of the Printed Manuscript

18. 30 copies of the respective paper are reprinted and distributed to the authors. The authors must cover any costs incurred by author's requests for printing more than 30 copies.

Release on the Internet

19. The papers published in the evaluation study are released on the internet with approval from the authors.

Copyright

20. Copyright of papers which appear in the evaluation study is attributed to the respective authors. Editorial right is attributed to the evaluation society.

Office

21. The office is in charge of administrative works for editing, distribution, and accounting.

Information for Contributors
(For English Papers)

Last revised on 20th July 2017

1. "The Japanese Journal of Evaluation Studies" is the publication for reviews, articles, research notes, and reports relating to evaluation.

2. "The Japanese Journal of Evaluation Studies" is primary published to provide opportunities for members of the Japan Evaluation Society (hereinafter referred to as "members") to exchange findings, and to contribute to further development of the study of evaluation both domestically and internationally. As a principle, this journal publishes the contributions submitted by the members. With the exception of requested papers, the first author must be a member. A submission (as the first author) is limited to one manuscript that has not been published or submitted in any form for another journal of academic association etcetera.

3. Adoption judgments of the manuscript are made at the discretion of the editorial board. Comments from two referee readers who are appointed for every manuscript are referred to in the screening process (the editorial board requests referee readers without notifying the author of manuscript).

4. Payment for the manuscript is not provided.

5. Papers published in "The Japanese Journal of Evaluation Studies" are released on the Internet at homepage of this academic society.

6. Regarding submission, manuscripts must be identified as one of the following categories: 1) article, 2) review, 3) research note, 4) report, and 5) others. However, the final decision of the category is made by the editorial board.
   "Article" is considered as a significant academic contribution to the theoretical development of evaluation or understanding of evaluation practice.
   "Review" is a paper which provides an overview of evaluation theory or practice.
   "Research note" is a discussion equivalent to the intermediate outputs of a theoretical or empirical study in the process of producing an "article".
   "Report" is the study report related to a practical evaluation project or evaluation.
   "Others" are manuscripts for special editions requested by the editing committee.

7. Manuscript Submission
   (1) Manuscripts may be written in either Japanese or English.
   (2) Correction by the author is only for the first correction.
   (3) English manuscripts should be submitted only after the English has been checked by a native speaker.
   (4) Submit manuscripts via email. Contact information including mailing address, telephone number, fax number, e-mail address, and the category of the manuscript should be clearly stated.
   For approved manuscripts, after necessary rewriting, the author needs to submit the final paper via email. Original figures, charts, and maps should be provided.
   (5) Total printed pages should not exceed 14 pages. Any cost incurred by printing more than 14 pages must be covered by the author.
   (6) The layout for English papers should be 30 mm of margin at left and right side, 10pt for font size, 43 lines on A4 paper (about 500 words per page). An abstract of 150 words should be attached to the

front. 14 pages are equivalent to 7,000 words but the body should not exceed 6,000 words to allow for the title, header, figure, chart, footnotes, and references. Please note that the number of pages may be more than expected depending on the number of figures included.

8. Mailing address

Office of Japan Evaluation Society at International Development Center of Japan
Shinagawa Crystal Square 12th Floor, 1-6-41 Konan, Minato-ku, Tokyo,
108-0075, Japan
E-mail: jes.info@evaluationjp.org

Writing Manual of the Japanese Journal of Evaluation Studies (For English Papers)

Revised on 18th September 2002

1. Text, Charts, Figures, Graphs, Diagrams, Notes, and References

(1) The paper should be written in the follow order:

First page: Title; the author,s name; Affiliation; E-mail address; Abstract (150 words); Keywords (5 words)

Second page: The main text; acknowledgement; notes; references

(2) Section of the text should be as follow:
1.
1.1
1.1.1
1.1.2

(3) Source of the charts, figures, graphs, and diagrams should be clarified. Submitted charts and others will be pzhotoengraved, therefore it is important that the original chart is clear. Pictures shall be treated as figures.

Figure 1 Number of Students in the State of ○○

Note:
Source:

Table 1 Number of Accidents in the State of ○○

Note:
Source:

(4) Citation of literature in the text should be, (Abe 1995, p.36) or (Abe 1995).
(5) Note in the text should be, (------.[1])

(6) Note and references should be written all together in the end.
Note
1 --------.
2 --------.

(7) Reference should list the literature in alphabet order, and arranged in chronological order. Follow the examples:

Book: author (year of publication). *Title of the book.* Published location: publishing house.

(e.g.) Rossi, P. H. (1999). *Evaluation: A Systematic Approach 6th edition*. Beverly Hills, Calif: Sage Publication.

Article from magazine: author (year of publication). Title. *Title of the magazine,* volume (number), page-page.

(e.g.) Rossi, P. H. (1999). Measuring social judgments. *American Journal of Evaluation,* 15(2), 35-37.

Article in Book: author (year of publication). Title. In editor (Eds.), *Title of the book.* Published location: publishing house, page-page.

(e.g.) DeMaio, T. J., and Rothgeb, J. M. (1996). Cognitive interviewing techniques: In the lab and in the field. In N. Schwarz & S. Sudman (Eds.), *Answering questions: Methodology for determining cognitive and communicative processes in survey research.* San Fransisco, Calif: Jossey-Bass, 177-196.

Book by two authors: surname, first name, and surname, first name. (year of publication). *Title of the book.* Published location: publishing house.

(e.g.) Peters, T., and Waterman, R. (1982). *In Search of Excellence: Lessons from America,s Best Run Companies.* New York: Harper & Row.

Book by more than three authors: surname, first name, surname, first name, and surname, first name. (year of publication). *Title of the book.* Published location: publishing house.

(e.g.) Morley, E., Bryant, S. P., and Hatry, H. P. (2000). *Comparative Performance Measurement.* Washignton: Urban Institute.

(note 1) If some references are from the same author with the same publication year, differentiate by adding a,b,c as (1999a), (1999b).

(note 2) If the reference is more than a single line, each line from the second should be indented by three spaces.

(e.g.) DeMaio, T. J., and Rothgeb, J. M. (1996). Cognitive interviewing techniques: In the lab and in the field. In N. Schwarz & S. Sudman (Eds.), *Answering questions: Methodology for determining cognitive and communicative processes in survey research.* San Fransisco, Calif: Jossey-Bass, 177-196.

Referee-Reading Guideline

The Japanese Journal of Evaluation Studies Editorial Board,
The Japan Evaluation Society
Approved on 10th September 2005

1. Content of the Referee-Reading Guideline

This Referee-Reading Guideline is to provide explanation of the main publication judgment, procedure of the referee-reading, to the members who submit the manuscript and for the members who are requested to conduct referee-reading in order to carry out the procedure efficiently and effectively.

2. Purpose of Referee-Reading and the Responsibility of the Author

Referee-reading is necessary for the editorial board to make decisions of whether submitted manuscripts are appropriate to publish in the Japanese Journal of Evaluation Studies or not.

If there is doubt or obscurity identified in manuscripts during the referee-reading corrections may be required. Therefore, referee-reading also contributes to the improvement of the submitted manuscripts. However, although the manuscripts are requested corrections, the author is still solely responsible in regards to the contents and it is not attributed to the referee-readers.

Referee-readers are two persons who are requested by the editorial board depending on the specialty or the field of the submitted manuscript. People who are not members of this academic society also may be requested.

3. Items of Consideration in Referee-Reading

Five points are considered in referee-reading, however, the importance of each may be different depending on the type of manuscript.

(1) Importance and utility of the theme
(2) Originality of the study
(3) Structure of the logic
(4) Validity of verification and methodology
(5) Contribution to evaluation theory and practice

- For the article, all of above five are considered.
- For the research note, especially (1), (2), (3), and (4) are considered.
- For the report, especially (1), (3), and (5) are considered.
- For the review, especially (3) and (5) are considered.

4. Attentions in submission of manuscript

Besides above five viewpoints, basic completeness as a paper is also considered, for example;

- appearance of the paper is organized
- written according to the writing manual
- described simply and distinctive
- verification data is appropriately used
- notes and references are corresponding with the text
- terminology is appropriately used
- no wording and grammatical mistakes
- no errors and omission
- no punctuation mistakes
- expression in English abstract is appropriate
- word count is according to the manual

The above mentioned forms and contents are also considered. There have been cases in which graduate students and practitioners posted without organizing the manuscripts as a paper. On those occasions, referee-reading was not conducted. Necessary consultation is strongly recommended prior to submission.

5. Judgment Cases in Referee-Reading

(1) In the case of the manuscript which is considered acceptable for the publication but is not yet complete:
The referee reader should evaluate carefully whether the paper can contribute to the development of evaluation theory or evaluation studies.

- Verification is lacking but the theory and formulation are useful for academic development.
- Analysis lacking but useful for formation and promotion of new theory.
- The literature review is not of a high standard but, the overall study is meaningful.
- Comparative study is not up to standard but is meaningful as an example of application.
- Analysis is lacking but it is meaningful as an evaluation of socially and historically important cases.
- Analysis is lacking but it is meaningful as an evaluation of particular social activities.
- Organization and expression are not up to standard as a paper but the contents are worthy to evaluate.
- Logic is not strong enough but useful in practice.
- The paper has significance as a report.

(2) In case of the manuscript which is considered as difficult for publication:

- Awareness of the issue or setting of the problem is indecisive.
- Understanding or analytical framework of notion of basic terminology is indecisive or inappropriate.
- There is a lack in credibility of data for the grounds of an argument.
- There is no clear point of an argument or appropriateness of proof.
- Organization of the paper and presentation (terminology, citation, chart, etc) are inappropriate (or not consistent).

6. Judgment

The final decision will be made on publication at the standing editors committee following one of four patterns (listed below). However, these judgments are not based on the number of errors but on the strength of the overall report. In the case of (3) and (4), there is a possibility to be published as a different type of paper. If it is published as a different type of paper, major rewrite concerning the number of words may be required.

(1) The paper will be published.
(2) The paper will be published with minor rewrite.
(3) The paper will be published with major rewrite, however as a different type of paper (review, article, research note, or report).
(4) The paper will not be published; however there is the possibility that it will be published as a different type of paper (review, article, research note, or report).

[KY1]

Japanese Journal of Evaluation Studies, Vol. 18, No. 2, June, 2018

Published by The Japan Evaluation Society
c/o International Development Center of Japan,
Shinagawa Crystal Square 12th Floor,
1-6-41, Konan, Minato-ku,
Tokyo, 108-0075, Japan
E-mail: jes.info@evaluationjp.org

Printed by Kenkosha, Co., Ltd.

ISSN 1346-6151

Japanese Journal of Evaluation Studies

Japan Evaluation Society