Issues of Program Evaluation in Social Work Education: A Review from the Perspectives of Evaluation Framework

Kenichi Shimura

Introduction

The purpose of this paper is to discuss program evaluation from several perspectives, and discuss the use of evaluation in social work education in Japan. Evaluating students' performance or skills in helping professions are complicated tasks, and educators in schools and supervisors in social work fields have been collaborating to create better system.

The Ministry of Health, Welfare and Labor introduced the standardized curriculum for social work education in 1988. The curriculum includes 15 course works, hands-on small group seminars and four-week-long social work practicum. Paper and pencil test is widely used to measure students' levels of understanding for those 15 course works; however, measuring the performance and skills for the seminars and practicum are still controversial. Social work schools and colleges modified the original curriculum based on the students' interests and the schools' characters within the range of the standard, and it made the issues more complicated. Japanese Association of Social Work School introduced a standard evaluation form for the practicum, but it also has been controversial, and not all schools use it. In this study, after discussing program evaluation in general, evaluation in social work will be discussed.

Part 1. Overview of Program Evaluation

Evaluation

As a social work and/or an educational profession, or as a scholarly practitioner, evaluation is considered a part of daily routine, either evaluating someone or something, or being evaluated. In social work process, which is known as the APIE cycle, evaluation is the final stage, and it comes after assessment, planning and implementation. The purpose of evaluation in this cycle is to evaluate how the social worker planned the program and how s/he implemented the planned program. S/he will decide if s/he would continue, modify or terminate the program based on the result of the evaluation. It will be recorded accordingly, and will be used as information for future planning. This cycle applies to social work education as well; however, it is questionable that if the evaluation is conducted properly, or if the result of the evaluation is used efficiently. Evaluation might have been misused or not been understood.

The following is several definitions of evaluation:

> Evaluation is the systematic collection of information to assist in decision making. One broad area of evaluation, referred to as program evaluation, provides systematic information about any of a variety of human services (Gredler, 1996).

In education, program evaluation refers to the set of activities involved in collecting information about the operation and effects of policies, programs, curricula, courses, and educational software and other instructional materials. Of importance is that program evaluation should not be confused with other forms of inquiry or data collection that are conducted for different purposes (Gredler, 1996).

Program evaluation is the systematic collection of information about the activities, characteristics, and outcomes of programs to make judgments about the program, improve program effectiveness, and/or inform decisions about future programming (Patton, 1997).

Fourth generation evaluation is a form of evaluation in which the claims, concerns, and issues of stakeholders serve as organizational foci (basis for determining what information is needed), that is implemented within the methodological precepts of the constructivist inquiry paradigm (Guba & Lincoln, 1989).

Evaluation vs. research

In the field of academics, people use the term "research", but in contrast, people in the practice field use the term "evaluation" although they both collect data and analyze them. Evaluation and research are similar in many ways; however, Gredler (1996) explained that there are four differences.

1. First, the major purpose of educational research is to test principles or theories that may be generalizable across space and time.
2. A second characteristic of educational research is that the researcher, depending on his/her particular area, determines the nature of the programs to be investigated.
3. A third characteristic of research is that methods and procedures are implemented so that individual values or preferences do not influence the outcome.
4. Fourth, the primary audience for the research is often researchers and theorists in a particular area of inquiry.

Patton (1997) described that program evaluation differs fundamentally from research in the purpose of data collection and standards for judging quality. Basic scientific research is undertaken to discover new knowledge, test theories, establish truth, and generalize across time and space. Program evaluation is undertaken to inform decisions, clarify options, identify improvements, and provide information about programs and policies within contextual boundaries of time, place, values, and politics. These two authors attempted to differentiate evaluation from research; however, some form of research also seek information to make decision rather than finding scientific truth like action oriented research. Thus it seems meaningless to differentiate two. The name, research or evaluation should be chosen when appropriate.

Another discussion is that evaluation is always deductive, because it is based on actions people worked intentionally to produce expected results. In contrast, research is not always deductive. Inductive discovery approach, such as grounded theory may not produce the result the evaluator needs to make a decision, because data is the only one that tells what it is and what to do next in grounded theory.

The birth of program evaluation

In 1932, thirty schools developed innovative curricula to serve student who couldn't find work. Effectiveness of these innovative curricula was assessed by students' learning and follow-ups after they graduated from high school; however, the traditional testing method couldn't measure these innovative curricula. The evaluation should be based on the important learning experiences that

are selected by the schools and the teachers (Gredler, 1996).　This study was called the Eight-Year Study, and it is considered as the beginning stage of the modern program evaluation.　Evaluation staff and teachers realized that the goal of the program included beyond knowledge and skills, and those were critical thinking and critical interpretation.　They made up innovative ways of assessing these goals (Gredler, 1996).

Gredler (1996) discussed that Tyler established three broad principles of curriculum evaluation from the Eight-Year Study.

1. Activities of measurement differ from evaluation.
2. A single term or a single dimension cannot describe a person.
3. Curriculum development is related evaluation.

The decade of 1960s in the United States was socially unrest; the assassination of JFK, Vietnam War and/or many civil rights movements, however the Great Society legislation was based on the belief that educational and social problems could be solved through a process of social engineering (Gredler, 1996).

The Elementary and Secondary Education Act (ESEA) was enacted in 1965, and it mandated evaluation and project reporting.　The early evaluation was conducted based on the Newtonian science paradigm; input-output concept and production or goal oriented concept; however, as society started to realize that this concept had the limitation to evaluate, study or explain social and human science.　Paradigm of science has been changing from Newtonian mechanical science to chaos theory or complexity, and these three discussions also probe that the evaluation principle follows the paradigm change too.　Thus, the area of evaluation grew wider and has been developed in many directions.　In the next section, several approaches will be discussed.

Various approaches to evaluate

There are as many evaluations as research methodologies; however, three evaluation theories are more superior that others.　Those are theory-driven evaluation, utilization-focus evaluation and fourth generation evaluation.

Theory-driven evaluation

Main theorists of theory-driven evaluation are Rossi and Chen, and Chen published the book "Theory-driven evaluation" in 1990.　Chen meant theory is generally defined as a set of interrelated assumptions, principles, and/or propositions to explain or guide social actions.　Chen pointed out that theory in program evaluation has been a neglected issue, and evaluation has been conducted by utilizing research method and traditions.　He continued that currently, there is a new movement to shift program evaluation from method-oriented evaluation to theory-oriented evaluations (1990).　Rossi and Chen argued that method-driven evaluations tend to maximize one type of validity at the expense others, and to avoid this problem, they pointed out the importance of program theory in simultaneously dealing with various types of validity.

Utilization focused evaluation

It is unfortunate that the results of research are seen or used by limited people, or only the members of the committee read someone's dissertation; however, results or reports of evaluation should be seen or read by decision makers and stakeholders, or it is a waist.　Patton (1997) explains

that utilization focused evaluation begins with the premise that evaluations should be judged by their utility and actual use; therefore, evaluators should facilitate the evaluation process and design any evaluation with careful consideration of how everything that is done, from beginning to end, will affect use. He continues that the focus in utilization-focused evaluation is on intended use by intended users. The followings are the criteria for utilization focused evaluation questions by Patton.

1. Data can be brought to bear on the question; that is, it is truly an empirical question.
2. There is more than one possible answer to the question; that is, the answer is not predetermined by the phrasing of the question.
3. The primary intended users want information to help answer the question. They care about the answer to the question.
4. The primary users what to answer the question for themselves, not just for someone else.
5. The intended users can indicate how they would use the answer to the question; that is, they can specify the relevance of an answer to the question for future action.

Utilization focused evaluation emphasizes that evaluations are for intended users to use.

Fourth generation evaluation

Fourth generation evaluation is a form of evaluation in which the claims, concerns, and issues of stakeholders serve as organizational foci (basis for determining what information is needed), that is implemented within the methodological precepts of the constructivist inquiry paradigm (Guba & Lincoln, 1989). According to them, constructive methodology is the approach that we propose as a replacement for the scientific mode that has characterized virtually all evaluation carried out in this century. It rests in a belief system that is virtually opposite to that of science; a kind of belief system that is often referred to as a paradigm. Thus, fourth generation evaluation is in the new science paradigm that is different from Newtonian and Cartesian mechanical worldview.

As definition includes, stakeholders play an important role as much as evaluators. Guba and Lincoln explained that why use stakeholders' claims, concerns, and issues as organizers. The followings are the answers.

1. Stakeholders are groups at risk.
2. Stakeholders are open to exploitation, disempowerment, and disenfranchisement.
3. Stakeholders are users of evaluation information.
4. Stakeholders are in a position to broaden the range of evaluative inquiry to the great benefit of the hermeneutic/dialectic process.
5. Stakeholders are mutually educated by the fourth generation process.

When evaluation takes place, stakeholders may be evaluated, use the information, and/or informant to the evaluation. The fourth generation evaluation bring them into the evaluation process, and attempt to produce the information that would be used to make future decision.

The fourth generation evaluator is responsible for the following:

1. Identifying the full array of stakeholders who ate at risk in the projected evaluation.
2. Eliciting from each stakeholder group their constructions about the evaluand and the range of claims, concerns, and issues they with to raise in relation to it.
3. Providing a context and a methodology (hermeneutic/dialectic) through which different constructions, and different claims, concerns, and issues, can be understood, critiqued, and taken into account.
4. Generating consensus with respect to as many constructions, and their related claims, concerns, and issues as possible.

5. Preparing an agenda for negotiation on items about which there is no , or incomplete, consensus.
6. Collecting and providing the information called for in the agenda for negotiation.
7. Establishing and mediating a forum of stakeholder representatives in which negotiation can take place.
8. Developing a report, probably several reports, that communicate to each stakeholder group any consensus on constructions and any resolutions regarding the claims, concerns, and issues that they have raised (as well as regarding those raised by other groups that appear relevant to that group).
9. Recycling the evaluation once again to take up still unresolved constructions and their attendant claims, concerns, and issues (Guba & Lincoln, 1989).

Standards for evaluation and guiding principles for evaluators

Two professional organizations, Evaluation Research Society and Evaluation Network merged in 1984 to form a new professional organization for evaluation, American Evaluation Association (AEA).　AEA is an international professional association of evaluators devoted to the application and exploration of program evaluation, personnel evaluation, technology, and many other forms of evaluation.　Evaluation involves assessing the strengths and weaknesses of programs, policies, personnel, products, and organizations to improve their effectiveness.　AEA's mission is to:

Improve evaluation practices and methods
Increase evaluation use
Promote evaluation as a profession and
Support the contribution of evaluation to the generation of theory and knowledge about effective human action.

Evaluation as a profession needed to share the standard to establish accountabilities and public awareness, and AEA announced and published the standards for evaluation in 1994.

Utility
The Utility Standards are intended to ensure that an evaluation will serve the practical information needs to intended users.
Feasibility
The Feasibility Standards are intended to ensure that an evaluation will be realistic, prudent, diplomatic, and frugal.
Propriety
The Propriety Standards are intended to ensure that an evaluation will be conducted legally, ethically, and with due regard for the welfare of those involved in the evaluation, as well as those affected by its results.
Accuracy
The Accuracy Standards are intended to ensure that an evaluation will reveal and convey technically adequate information about the features that determine worth or merit of the program being evaluated.

Patton (1997) stated that implementation of a utility-focused, feasibility-conscious, propriety-oriented, and accuracy-based evaluation requires situational responsiveness, methodological flexibility, multiple evaluator roles, political sophistication, and substantial doses of creativity, all element of utilization-focused evaluation.　In addition to the Standard, AEA developed the guiding principles for evaluators as professionals.

Systematic Inquiry
Evaluators conduct systematic, data-based inquires about what is being evaluated.
Competence
Evaluators provide competent performance to stakeholders.

Integrity/Honesty
Evaluators ensure the honesty and integrity of the entire evaluation process.
Respect for People
Evaluators respect the security, dignity, and self-worth of the respondents, program participants, clients and other stakeholders with whom they interact.
Responsibilities for General and Public Welfare
Evaluators articulate and take into account the diversity of interests and values that may be related to the general and public welfare.

AEA has about 3000 members in all 50 states in the US and as well as over 50 foreign countries. Evaluation became a profession and an area of interdisciplinary social science worldwide.

Part 2. Evaluation in Social Work

Research and evaluation in social work

Social work profession has always had an ambivalent relationship with research (Anastas, 1994). The profession was born based on and still strongly supported by religious belief, thus the profession has been sensitive to be scientific; however, it was also important to prove the effectiveness of the service through research to establish professional accountability among other professions.

There is no question that research and evaluation play very important part of social work today. For social work academics, conducting research to test existing theories and to produce a new knowledge for the field is in their job descriptions. For social workers, supervisors and administrative staffs request evaluating the case, the program and community resources. Council on social work education has the standard curriculum for accredited schools, and the followings are the research content of the standard.

Curriculum content for BSW
- The research curriculum must provide an understanding and appreciation of a scientific, analytic approach to building knowledge for practice and to evaluating service delivery in all areas of practice. Ethical standards of scientific inquiry must be included in the research content.
- The research content must include quantitative and qualitative research methodologies; analysis of data, including statistics procedures; systematic evaluation of practice; analysis and evaluation of theoretical bases, research questions, methodologies, statistical procedures, and conclusions of research reports; and relevant technological advances.
- Each program must identify how the research curriculum contributes to the student's use of scientific knowledge for practice. (Council of Social Work Education, 1994)

Curriculum content for MSW
- The foundation research curriculum must provide an understanding and appreciation of a scientific, analytic approach to building knowledge for practice and for evaluating service delivery in all areas of practice. Ethical standards of scientific inquiry must be included in the research content.
- The research content must include qualitative and quantitative research methodologies; analysts of data, including statistical procedures; systematic evaluation of practice; analysis and evaluation of theoretical bases, research questions, methodologies, statistical procedures, and conclusions of research reports; and relevant technological advances.
- Each program must identify how the research curriculum contributes to the student's use of scientific knowledge for practice. (Council of Social Work Education, 1994)

Not only social work academics, social work practitioners are expected to have knowledge in research

and evaluation to guide and support their interventions.

In the field of social work, research is defined that a structured inquiry that utilizes acceptable methodology (i.e., quantitative and qualitative) to solve human problems and creates new knowledge that is generally applicable (Grinnell, 2001). Reamer (1998) summarized the four main functions of social work research and evaluation as follows.

1. To assess the effectiveness of interventions
2. To conduct needs assessment
3. To draw on empirically based literature and findings to inform and provide guidelines for practice
4. To create and disseminate knowledge for use by other professionals

When looking at the main functions of research and evaluation, they are closely related or parallel to the process of social work intervention, and they are not distinguishable from the functions, however Grinnell (2001) explained that research studies fall into the four broad categories: (1) descriptive, (2) exploratory, (3) explanatory, and (4) evaluative. Thus, evaluation is an area of research in social work. He explains that evaluation research differs form other forms of explanatory research because evaluation research considers the implementation and effects of social policies and programs. Reamer (1998) also uses the same four categories as social workers main goal, but when he classifies the types of research he put explanatory and evaluative category into one. Simmons explained that the social science research fall into the four matrixes. The author modified the matrix.

Social research matrix

	Generating theory/hypothesis 〈Inductive〉	Verifying theory/hypothesis 〈Deductive〉
Descriptive	Ethnography * Participant observation * Interviews * Human document	Descriptive statistics * Questionnaire (Survey) * Existing statistics
Explanatory	KJ Method, Grounded theory * Participant observation * Interviews * Human document	Inferential and empirical statistics * Questionnaire (Survey) * Quasi-experimental design * Experimental design

When using this matrix, evaluation fits into under explanatory categories. When evaluators want to know what is going on in the particular field, it is inductive study, and when evaluators wants to find out if the program is effective, it is deductive study.

Another way of classifying research and evaluation is pure research and applied research. Pure research means generating or developing theories, and applied research is hypothesis testing or verifying the existing theories. In this case, evaluation fits into applied research.

Program evaluation in social work

Applied research in social work usually focuses on clinical work, needs assessment, or program

evaluation (Reamer, 1998). He continues that when conducting program evaluation in social work, social workers need to address four major issues.

1. Who wants the evaluation and why?
2. What are the goals of the program evaluation?
3. How will the results be used?
4. Who should design and carry out the evaluation?

In addition, he recommends including stakeholders, and this basically follows general rules of evaluation. In the next section, the evaluation of social work education in Japan will be discussed by utilizing these considerations.

Evaluation of social work education

When discussing this issue, there are two areas to be addressed: (1) social work education, (2) students' progress. Social work education includes the curriculum or program evaluation and the instructors' teaching performance, and the students' progress includes knowledge, skills and their attitudes toward the profession and the clients. Both areas are interrelated, and thus it is complicated to evaluate.

(1) Social work education

Who wants the evaluation and why?

Professional organizations have the biggest interests evaluating on social work education. There are four nation wide professional organizations for social work professionals, and two of these are for social work schools and the others are for social workers. Japanese Association of Schools of Social Work (JASSW) and Japanese Association of Schools of Certified Social Worker (JASCSW) are the organizations for social work schools. Four organizations have been working closely to promote social work and the education.

JASSW was established in 1955 with 14 social work schools, and has been evaluating prospective schools to join the organization to keep the minimum standard to claim a social work schools. JASCSW was established in 2001, and is growing rapidly, because social work schools for certified social workers are eligible to join automatically. The ministry of health, welfare and labor initiated to establish this organization, because it is useful to communicate with one organization than over two hundred schools individually. Certified social worker is a national certification that the ministry is responsible to supervise the profession and the level of education.

Professional social workers' organizations are also interested in the area of its education. They share the responsibility especially the fieldwork experience, internship and practicum work. These are important parts of the education.

What are the goals of the program evaluation?

The main goal of the program evaluation is to provide the best possible service to the clients and to protect them. To accomplish the goal, the knowledge and skills of social workers and the professional's education and training should be evaluated. Thus, goals of the program evaluation in social work education would be to maintain the standard of social work education.

How will the results be used?

The following is a case of the evaluation to join the JASSW. Basically this evaluates if the

school meets the standard to claim as a social work school or not, thus the results will be used to make the judgment.

Who should design and carry out the evaluation?

This is the same case as above. The board members make the judgment according to the standards that was set by the association. When the association was established in 1955, there were only 14 social work schools, however the population of older adults grew, and the normalization principle was introduced, and the advancement of technology made some people with disability independent in many aspects of their lives, thus the needs of social workers grew as well. 168 schools are members of JASSW as of April 2002. The annual membership meeting is the highest decision-making body, and committees and subcommittees propose and report various matters. Standard committee is one of those committees, and it designs and carries out the evaluation.

(2) Students' progress

Who wants the evaluation and why?

There are many situations, courses and events to be evaluated. Professors need evaluations to grade courses, and students want to know how they are doing. Parents as sponsors want to know if they are spending their money wisely, and the results of the evaluation helps to know that. When the chairperson of a social work institution wants to hire a new graduate, s/he may request students' grade reports. From a professor's perspective, evaluations of students' progress are for mostly grading sake, however the results also tell how we taught.

When discussing field-related courses, such as practicum and internship, not only professors evaluate the program but also supervisors in the field. The reason they evaluate the course is to know if students are learning skills and knowledge to be a social worker, and their levels of mastering and understanding. Evaluation is necessary to make them as reliable social workers to serve clients and communities.

What are the goals of the program evaluation?

The goals of the program evaluation of social work education are to know if students are learning to be social workers. The goals include how much students learn knowledge and master skills.

How will the results be used?

Basically the results would be used to grade the students. In addition, the results are always useful to evaluate how professors design and taught the class or the course. If they become reliable social workers, the delivery system of social work education is good, but if not they should evaluate the curriculum itself.

Who should design and carry out the evaluation?

Usually professors design the evaluation and s/he and the field instructor carry out the evaluation. Unfortunately professors evaluate the students, but not many of them evaluate the course itself.

Problems of program evaluation in social work education in Japan

As described above, the program evaluation in social work education in Japan is chaotic. There

are several reasons that create the situation. First, although when joining, professional organizations evaluate the curriculums and professors professional activities, they don't have authorities to monitor the membership schools. Even they can evaluate the program; it will be a time and money-consuming project or they have to mandate to submit self-evaluation to the association, but can we really relay on the report? The ministry of education, culture, sports and science recommend colleges and universities to conduct periodical self-evaluation, however most of them are garbage.

Second, no one evaluates each course except periodical self-evaluation, but as described above, it is not powerful evaluation. Professors distribute anonymous evaluation at the end of the semester, it is not systematically done, and/or students don't care much, because they completed the course already. Thus the evaluation of the students' progress is currently the only options to evaluate the course. If many students are learning new stuff and mastering new skills well, the professor could assume the course went well.

Then third problem emerges: how professors evaluate the students' progress especially the areas of caring attitudes, or students' passions to understand others in the field practicum work and the internship. These are necessary talent to be social workers. JASSW has created a standard evaluation form for the internship, but the evaluation form created new confusion, because it includes the areas of emotion, such as passionate. The evaluation form is not only tool to grade the internship course, and many professors now utilize the portfolio system to evaluate the students' progress.

Conclusion

Program evaluation is a developmental area of Japanese higher education, because the paternalism still remains in professors-students relationships. Thus, professors teach what they want to teach and do not care much about what they need and learn in many areas; however, the population of 18 years old is decreasing, and higher education is facing the post golden age. Program evaluation in each course, curriculum of the department and the school will be necessary to establish and maintain accountabilities to stakeholders soon.

References

Anastas, J. W. (1994). *Research design for social work and the human services.* New York: Lexington.

Chen, Huey-Tsyh. (1990). *Theory-driven evaluations.* Newbury Park, CA: Sage.

Gredler, M. E. (1996). *Program evaluation.* Englewood Cliffs, New Jersey: Prentice Hall.

Grinnell, R. M., Jr. (2001). *Social work research and evaluation: Quantitative and qualitative approaches.* *6th edition.* Itasca, Il: F. E. Peacook.

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

Patton, M. W. (1997). *Utilization-focused evaluation: The new century text.* Thousand Oaks, CA: Sage.

Reamer, F. G. (1998). *Social work research and evaluation skills: A case-based, user-friendly approach.* New York: Columbia University Press.