Public Policy Evaluation using SD Group Modeling

Makoto Ikeda, Professor, Faculty of Regional Development Study, Toyo University
1-1-1, Izumino, Itakura, Oura-Gun, Gunma Prefecture, 374-0193, Japan
+81-276-82-9006(TEL)/+81-276-82-9802(FAX)
mikeda@toyonet.toyo.ac.jp

Toru Suetake, Researcher, Research and Development Initiative, Chuo University
Soshigaya-1-11-23, Setagaya-ku, Tokyo Japan, 157
+81-3-3266-7684(TEL)/+81-3-3266-7616(FAX)
tohru.suetake@jp.kpmg.com

Abstract

Evaluation of public policy and projects of government-to-government assistance are quite common in Japan. Evaluators also use logical model bases for evaluation. However, they continue to use simple tree type model without incorporating loop or feed back effects. The author has insisted that SD modeling is applicable for quantitative evaluation of public policy. In addition, the trend in evaluation is shifting to use a more quantitative model and evaluators try a system thinking modeling to make a clearer more logical structure of the evaluation target. Application of SD modeling is primarily for planning and quite a few are even being used for evaluation purposes outside Japan.

We submitted our discussion paper to the Japan Evaluation Society and conducted academic discussions on this issue in December 2004. Most evaluators believe their logic model is not perfect, however, they feel that the present ST/SD model is also too complicated. We also try to conduct experimental study and find traditional group modeling methodology is not applicable well for evaluation of public policy. In this paper, we wish to discuss opportunities of ST/SD modeling for public policy evaluation with our new methodology.

Keywords: Evaluation, evaluation model, public policy evaluation, and system thinking model

1. History of public policy evaluation in Japan

Seldom used SD: system dynamics model for the evaluation of public policies is not only used in Japan but also in Europe and North America. However, the trend in evaluation is shifting to use a more quantitative model and evaluators try to think like system thinking modeling to clearly make a logical and structured evaluation of the target.

Historical background for starting public policy evaluation is different in Japan and other Western countries. In Japan, there are mainly two areas for applied public policy evaluation namely, central and local government, and government-to-government assistance. Since long economic recession insists on budget cutting for public policies, and introduction of new concept of so called new public management, the Government of Japan and local governments need to explain efficiency and effectiveness of their public policy in term of accountability. Also, in government-to-government assistance area, so called ODA: Official Development Aids, since 2000 the Government of Japan could not submit and maintain the position of top donor any longer. In this area too, the Government of Japan and government assistance agencies need to explain efficiency and effectiveness of projects and program of assistance as accountability to taxpayers. Governments (municipal governments and central governments) selected consultants as professional evaluator and request them to conduct evaluation to their projects and programs (ex. infrastructure development project, government to government assistance projects). Government as purchaser use the results of the evaluation as third party review or auditors review and use as prove of their works conducted efficiently, effectively, sustainable way and their work provide impact to the society. In short, professional and independent evaluator evaluates and proves as works and activities of the government done properly to tax payer so called accountability.

It may seem odd that the Government of Japan stresses on accountability while other assistance agencies of Western countries concentrate more on lessons learned from evaluations. Trends in Japan also seem to be shifting to more on lessons learned, because accountability is regarded more of an audit issue than evaluation issue. Stress on accountability for evaluation is the result of poor audit system in Japan, especially audit of government organizations and agencies. In next section, we briefly explain

method applied by those Japanese evaluator use and problems they faced with.

2. Issues of evaluation using quantitative model

For public policy evaluation in government-to-government assistance area, normally, DAC: Development Assistance Committee evaluation principles and three kinds of techniques are used. This DAC's evaluation principles are somehow very reasonable. Since tacit objectives of policy evaluation in Japan aim to explain rationality of the policy, program and projects to taxpayer, DAC's evaluation principle is very useful for pretending the evaluation was done under some international standard.

2-1: DAC evaluation principles

DAC of OECD: Organization for Economic Cooperation and Development announced the Principles for Evaluation of Development Assistance in 1991. In this principle, DAC suggests evaluating the following 5 issues according to the most popular Japanese evaluation textbooks (e.g. FASID 1999, FASID 2000):

- Objectives Achievement
- Impact and Effects
- Efficiency
- Rationale
- Sustainability

Somehow this principle is treated as supreme law for evaluation of government-to-government assistance projects in Japan. Almost all project evaluation of government-to-government assistance is done using those 5 issues. However, in many cases, evaluators do not consider logic model, however, insist they conduct quantitative evaluation when they can measure some performance indicators. Quantitative evaluation on Efficiency is rather easy. In government-to-government assistance projects, project staff normally use PCM: Project Cycle Management methodology and compose a PDM: Project Data Matrix, as a summary sheet of the project plan. This PDM includes inputs and outputs. In the evaluation stage, evaluator can use this PDM. Efficiency can be calculated with inputs to the project and outputs from the project. Objective Achievement is also rather easy for measuring progress and comparing them with project targets. But arguments come mainly on how to evaluate Impact and Effects. Since there are two structure and program layers or policy composed with set of projects, it is quite difficult to measure impact and effect to outcome of a program (upper layer) from result of another project (lower layer). Without using a structure model, it is impossible to measure the Impact and Effects of a project to a program or public policy. In PDM, they try to simplify this complicated relation between program level and project level. They set important assumption and if assumption is right as stand still, the overall goal (one of outcomes in program or public policy) can immediately be achieved. For the purpose of project evaluation, it works. However, for the purpose of program or public policy evaluation, relations between projects are always complicated and not simply like that. The arguments and discussions in the Japan Evaluation Society are mainly on difficulty of evaluating impacts and effects of project to program and public policy. For evaluating impacts and effects of project to program and public policy, they start to use a logical model. There does not seem to be general rules for quantitative evaluation on Ration-able and Sustainability. Sustainability contains both financial sustainability and non-financial sustainability such as organizational, social or cultural sustainability.

Though evaluation of public policy without government-to-government assistance do not follow DAC's principles, evaluation of impacts and effects between project level and program level is quite difficult and the same arguments and discussions are continued.

2-2. Evaluation methods

Evaluator begins by using logical model and based on the logical model, they will conduct evaluation by one of three methods: Score Evaluation, Balanced Scorecard or Statistic. However, logic adapted in the model is still straight thinking and not consider feedback thinking.

PDM: Project Design Matrix

Narrative summary	Objectively Verifiable Indicators	Means of Verification	Important Assumptions
Overall goal			
Project purpose			
Outputs			
Activities	Inputs		
			Preconditions

Evaluation using PDM

	Efficiency	Effectiveness	Impact	Relevance	Sustainability
Overall goal			↕	↕	↕
Project purpose		↕	↕	↕	
Outputs	↕	↕			
Inputs	↕				↕

Figure-1: PDM and evaluation using PDM

(1) Score evaluation

This technique is quite common and popular. If character of evaluation target can categorized into several groups, then select character for evaluation and set performance indicators that may perform the character. After setting the performance indicators, evaluator measures the performance indicators and compares with standards. Under normal circumstances, the evaluator can evaluate and examine results.

For example, take into consideration how school pupils are evaluated. Normally, the teacher conducts a series of tests in mathematics, science, social science and English to evaluate a pupil in primary school. If that pupil receives high scores for all subjects, the teacher concluded he or she is a good pupil. But if not, he or she must study harder to pass the next test. Honestly, it is difficult for me to explain to my children about the logic behind this evaluation, however, most teachers do it in such a way without having any doubt.

In general, consider a model composed with characters that can be measured by indicators and evaluated. Or, in other words, evaluator divides evaluation target into several categories and sets performance indicators to divided categories. If category and performance indicators keep mutual independence and total consistency of rules, then, evaluator can measure performance indicators, compare the result with standards, and make some judgment such as bad, ordinal or good.

Evaluator uses straight-forward-logic model as bases of this evaluation method as shown in Figure-2.

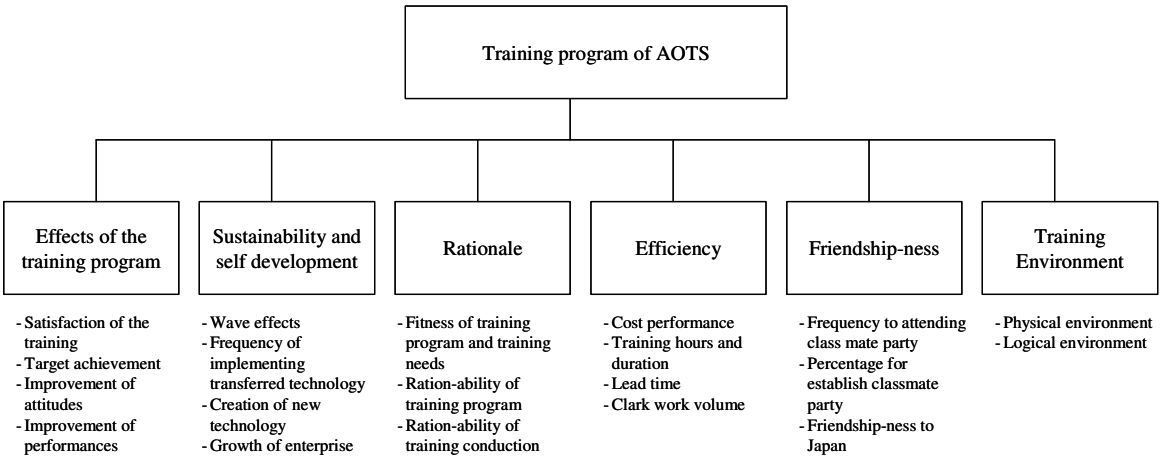


Figure-2: Sample score evaluation model of AOTS's training program (Itakura, 2002)

Again, such a model does not actually explain the deep characteristics of targets. It is just a list of what evaluator believes are characters of target and possible to evaluate based on measure or count of quantitative attribute with elements. Evaluator designs questionnaire sheet based on this model and conducts questionnaire survey. Then collects and calculates the answer. For example, questionnaire survey begins as "Are you satisfied with this training". Then collects answer sheet and tallies up number of "Yes". Under these circumstances, questions are being answered and counted up simultaneously. But if

attendee of the training course answers all questions pertaining to the training course in a positive manner, then doubt as to whether the training program is effective. I believe the evaluation should not be conducted in this manner. Instead, the evaluation should be conducted with a more total view, taking into consideration the background and environment of target. Under such circumstances an attendee should also be evaluated as to how effective the training course is after returning to their home country. Gathering feedback of this sort would prove to be an effective way to evaluate the training.

(2) Balanced Scorecard

Quite recently, evaluator began using Balanced Scorecard models for evaluation of public policy in Japan. There are also many discussions and arguments in the System Dynamics Society regarding the concept of Balanced Scorecards, which is now becoming popular.

Concept model of Balanced Scorecard looks more like the system thinking model and much more familiar to the SD practitioner. However, if the model is looked at closely, these models do not clearly explain feedback and just show the relation of elements. We believe the concept of Balanced Scorecard is very useful for clearly explaining the nature of evaluation target, however, it is not useful for conducting evaluation.

Moreover, starting from a balanced Scorecard model, the system thinking model is much easy to develop and much easier to evolve to the System Dynamics model for both qualitative and quantitative analysis. There are several trial models for quantitative evaluation witch are shown in figure 3. Many evaluators feel more comfortable with Balanced Scorecard charts rather than ST/SD loop diagrams, even with the lack of feedback loop and so many questions regarding detail relations. The problem with qualitative analysis used in Balanced Scorecard is that it is missing so many elements to rigorously determining relations. It may be an effective method to explain structure of a model as well as being accepted by experts because of module structure, however, more elements need to be added for conducting quantitative simulations.

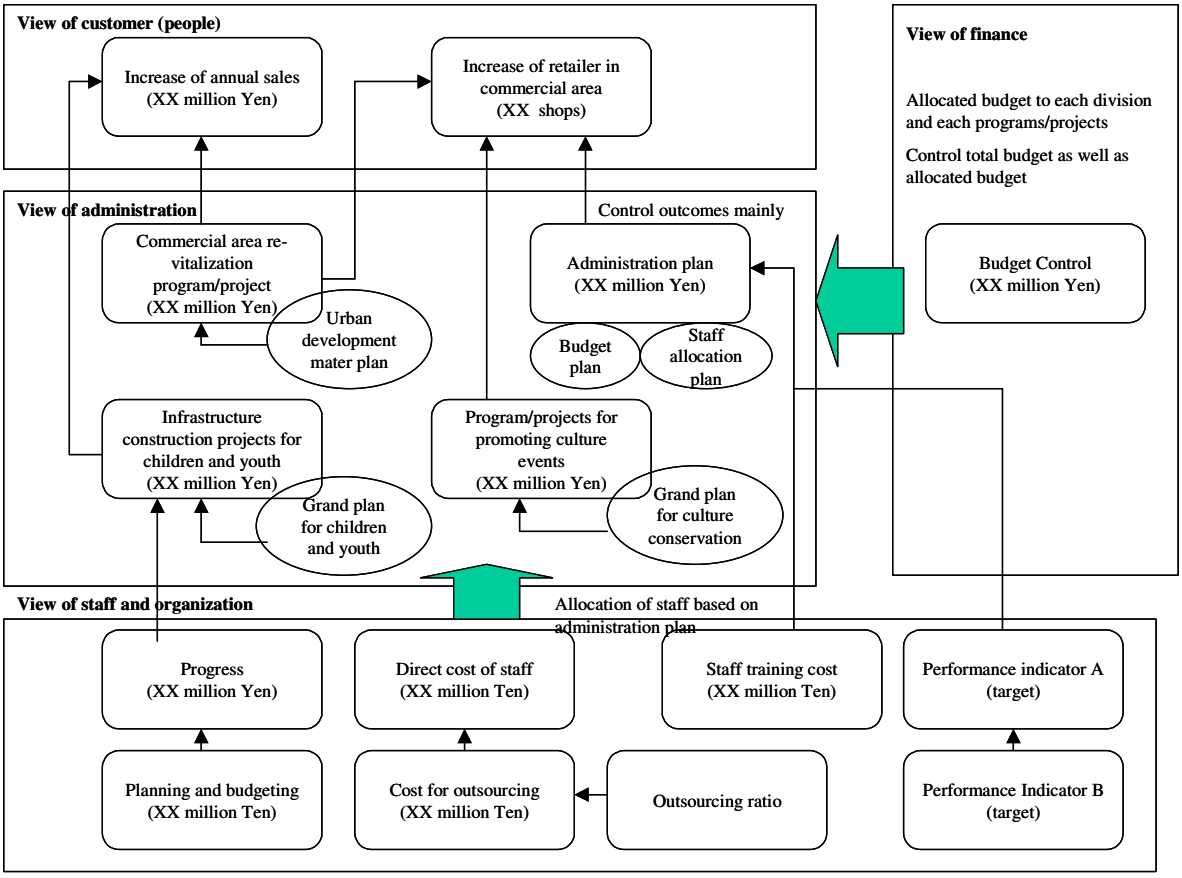


Figure-3: Sample of Balanced Scorecard to evaluate grand policy for revitalization of commercial area, by Morita (2004)

(3) Statistics technology

Using statistical technology for evaluation is actually the most popular method in Japan and

evaluators believe using statistics is the only scientific style for evaluation. Normally, they build models similar to Figure-2, using a straight forward thinking model, and then conduct questionnaire survey. Many times an interview survey is also conducted for a deeper understanding of background and to confirm the relation of elements, which then are analyzed to evaluate strength between elements with statistic techniques connecting the most ideal relations.

Figure-4 below shows statistic evaluation by Kishi (2004) designed to evaluation a water supply project in Indonesia with ADB: Asian Development Bank. To evaluate the effectiveness of water supply project, she first writes up project outcomes and tries to link them with the social impacts.
The study focus is to confirm through awareness the relation of direct project outcomes and indirect project outcomes such as easy access to water and proper water usage. She can prove existence of general relation between project outcomes and social impacts with statistic techniques; however, it does not try to determine quantitative relation between project outcome and social impacts such as percentage of reduced water borne diseases.

Every water supply projects primary objective is to improve hygiene conditions of people and contribute to socio-economic development by enforcing health condition of workers. However, in my experience, improving the relation of hygiene and water supply is rather easy but improving the relation with contribution of water supply to socio-economic development with straight-forward thinking model is difficult. Statistics are sometimes very useful and on occasion clearly illustrate relational strength. Nevertheless, statistics show only compound relation; therefore, we need a clearer model to identify what exactly we wish to obtain before conducting a statistic analysis.
We do not deny the validity of statistic techniques for evaluation; however, we do not believe statistics is the only way to attain a scientific evaluation that is widely accepted by evaluation practitioners.

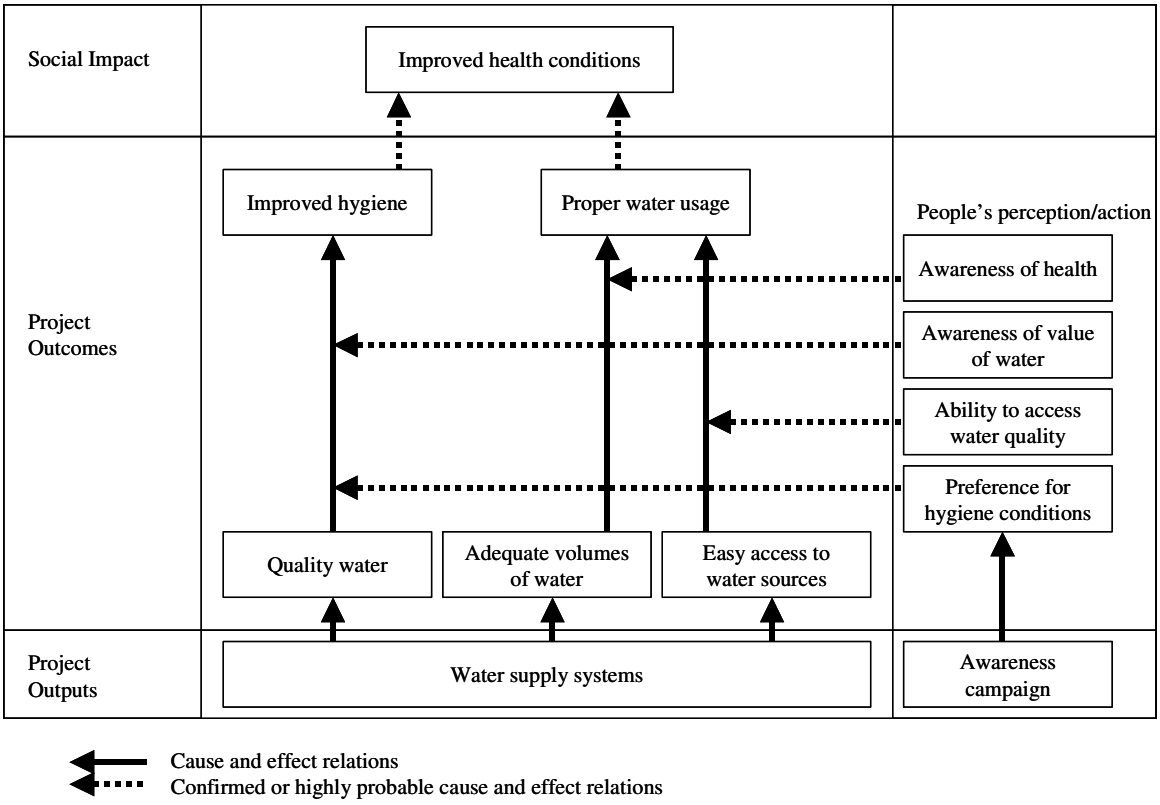


Figure-4: Sample logic model for statistic evaluation of water supply project by Kishi

2-3 Problem of evaluation

There are mainly two problems for evaluation applied with DAC's evaluation principles. Firstly, model used in evaluation is basically structured model or logical model like figure-2. It may no necessary to mentioning about problem and weakness of logical model. Luck of feedback loop, luck of time delay, etc.

Second problem is, evaluator looks very short ad narrow insight and skip off many external elements. This may makes trouble for difficulties to incorporate evaluation of outcomes. Outcome

means total results or effects that could be achieved with many projects and activities in this paper. Target project or program is only one of contributor for this outcome. Examples of outcome are, growth of local economy (GDP) and reduce of water born diseases in case of water supply project. In water supply project, municipal government could be supply safe water to people. But availability of safe water is only one of reason for reduce of water born diseases. Another factors such as improvement of healthcare system and public hygiene promotion program are also necessary for achieve this outcome.

As I already mentioned, DAC's evaluation principles included evaluation of impact and evaluation of rationale. Evaluator could not skip off to evaluating outcomes correctly for evaluate impact of the target project/program to the outcome.

SD/ST practitioner already knows well for feedback incorporated SD/ST model solved those problems. This issue is not our target for discussion in this paper. Rather we wish to discuss about how to evaluate target project or program with DAC's evaluation principles applied with SD/ST modeling.

Using ST/SD as evaluation model and conduct evaluation of the 5 issues is not so difficult. Some of them do not necessarily need a special model. For example, evaluation of "Objectives Achievement" may be rather easy and it may not be necessary to use a quantitative model in most cases. Of course using ST/SD model and conduct simulation is useful particularly in cases where target project does not achieved the objectives perfectly. It may be useful to confirm whether the objectives are truly achievable or not.

Efficiency is also a rather easy element of evaluation. Again, find out project budget and input of human resources, compare value analysis results of outputs, and divided with inputs. If project spends 100 million dollar and produces one billion values, it may be efficient. But if it only produces 100 million values, how could it be efficient? Again, actual evaluation is not so simple, but concept is simple.

Evaluation of sustainability does not depend on the model so much. Rather, it is considering the capability of a target organization and the environment. If target organization has sufficient capability for keeping operational activity in term of financial and technical skills, and socio-economic environment is stable and sustainable, the results of the project may have sustainability. However, it is not necessary to say ST/SD simulation is useful before making a conclusion of sustainability. If long-term simulation results show stability, it is truly sustainable. But impact and effect may need to be considered regarding the relation between factors used with model. Judgment of this issue basically needs some logic model. For this purpose, many evaluators make a model as previously shown. However, as mentioned beforehand, evaluator uses straight forward logic model. Discussion and argument in the National Conference of Japan Evaluation Society with our discussion paper focus on a different model used in evaluation. Evaluators insist that the straight forward logic model is the best and only practical model. Furthermore, our suggested feedback loop model seems too complicated for them. They also insist that the model should be as simple as possible. Most difficult issue is rationale. In Japan, the evaluator translates meaning of adequacy or proprieties and evaluates adequacy of target, objectives and outputs. If evaluated using this context, evaluator should list all alternatives and examine all of them. Of course no evaluator should try to challenge in such way. Incidentally, discussion of this paper focus on this issue and suggests that it may be possible with group modeling techniques. The following introduces how we conduct an evaluation.

However, it is one big difference with normal group model building in our approach. Group model building rather focus to develop one but best model for mainly find solution. For this purpose, make cross-organizational team and incorporate professional knowledge to make one but best model. Well, in such one but best model could very useful for evaluation of objectives achievement, impact and effects, and efficiency. But one model could not prove as target project or program was done rationale way. For solving this problem, we suggest to incorporate multiple evaluators and conduct evaluation on rationale with their SD/ST model as much possible alternatives. In short, we suggest evaluation with multi models instead of single one.

3. Case study for using modeling technology with system thinking

For development of evaluation method on public policy, specially evaluation of grand policy, we conduct experimental study with cooperation of students from Professor Makoto Ikeda's class as well as Tatebayashi Municipal Government. We already know the municipal government conducted survey for their public policy and has results with traditional questionnaire survey and analysis method. We

consider opportunity of public evaluation using SD/ST model but soon find single model is not sufficient for evaluation of rationale. Professor Ikeda has been teaching SD/ST modeling in his class and it is very good opportunity for conducting evaluation with multi models focus on evaluation of rationale.

(1) Basic concept *)

Using group ware via the Internet, we requested building a group model with the students of Professor Makoto Ikeda's class. Evaluator is also a student of his class. The evaluation target is the Tatebayashi Municipal Government grand policy. We ask the evaluator to build as much as possible and conduct simulations on their model. We do not insist that they create one model, but rather as many models as necessary to ideally execute the grand policy if he or she were mayor of the municipal government. Group modeling introduced in SD society is focused on building a single but best model, with discussion by participants. However, this time, to focus on evaluation, we supported technical issues only. Moreover, we used discussion for standardized basic elements of a model, but we do not insist on constructing one single model or creating limitations of modeling with those standardized basic elements. Logic between standardized basic elements is free for model builder. In addition, adding new elements to the model to realized his or her idea for simulation is free.

Under these circumstances, we collected all created models and conducted simulations. All possibilities evaluator could find to obtain conclusions on adequacy or proprieties of present grand policy of Tatebayashi Municipal Government were compared.

(2) Tatebayashi City

Tatebayashi City was already introduced in our study last year. This municipality located between two prefectures of Northern Tokyo, Tochigi and Gunma Prefecture. Population is around 80 thousands and the major industry is agriculture. They do not have a major industrial industry at all. However, they have a lot of attractive but not well-known tourism resources including cherry blossom in spring, azalea and summer festival that only local people know well. It is a typical local town in Japan.

Tatebayashi Municipal Government issues grand policy for development of the municipality. It is rather a vision and strategy paper of long term development until the year 2010. Based on this vision and strategy, they issue basic concepts. Again, based on this concept, they announced 6 basic policies, with two master plans already developed.

- Development Concept of Tatebayashi City

1) Basic Concept: Peaceful but vital green park civil society

2) Four concepts to achieve and support basic concept

2-1) Characteristic of municipal based on clear urban development plan: Aimed at developing attractive and vital city based on environmental, historical heritage, and cultural considerations in accordance with structure of city and traffic linkage system.

2-2) Safe and friendly society: All citizen can live safely in various life stages as well as feel a sense of satisfaction of his/her life in this city with support of health, medical and welfare system.

2-3) Safe and comfortable living with concern for the environment: Consider building a recycle society, and develop more infrastructure including sewerage system, public parks, and fire protection system for more comfortable living.

2-4) Vitalized city that is more adapted to the changing world: Create vitalization from re-development of commerce area, re-vitalization of commerce, re-vitalization of agriculture, development of new local industries in support of job creation.

3) Six basic policies for supporting four basic concepts

3-1) Infrastructure development: Realize characteristic of city based on clear urban development concept aimed at achieving peaceful but vital green park civil society.

3-2) Create a city of safe and comfortable living: In harmony with natural environment such as the beautiful night sky with thousands of twinkling stars and small fish swimming in rivers and lakes.

3-3) Improve social welfare system: Improving social welfare where people can feel safe, while comfortably living in support of each other.

3-4) Develop lifelong education society: Support lifelong education system to continually develop rich minds of human resources with positive temperament, capability and vitality for creating a rich cultural city.

3-5) Industry development: Developing local industry based on culture heritage and natural environment aimed at vitalizing the city while adapting to the changing world.
3-6) Implementing plans: Implementing changes with participation of citizens to create a new 21st century municipality.

In 2004, Tatebayashi Municipal Government conducted a questionnaire survey for evaluating their grand development plan. During this evaluation, they used typical statistic techniques to fulfill distributed questionnaires and requests. After which they collected and tallied the answers. Based on the number of answers, they assessed the grand policy. However, it made no account to uncover whether people were satisfied with the grand policy. We tried to identify all possible alternatives of grand polices among evaluators to discover which has the best and most accurate amount of alternatives.

(4) Process of group modeling for evaluation of grand policy
We guide evaluator to build models as shown in the following steps.

Step-1: Orientation.
Help evaluator understand background of grand policy, show them promotion video of Tatebayashi City as well as conducting group tours of discussions with members of the Committee of Re-Vitalization of Commercial Area of Tatebayashi City. Evaluator should understand socio-economic background of Tatebayashi City including historical heritage, tourism resources, activity of people and opinions of citizens.

Step-2: List up and standardization of keywords
List keyword candidates for model from evaluators using Internet. Keywords are basically performance indicators, should be nouns, variable, and not a negative sentence. Evaluator find candidates of keywords from the gland policy of Tatebayashi municipal government.
After collecting keyword candidates, use grouping techniques to combine and find one word describing more precisely the character of the group.

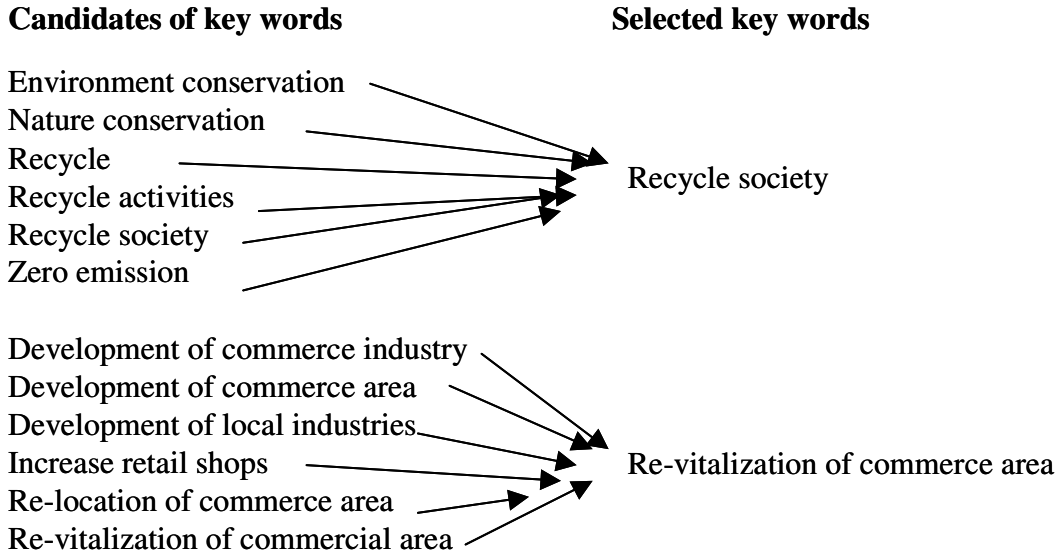


Figure-5: Grouping and selection of keywords

Step-3: Select 10-15 most important keywords:
Select 10-15 most important keywords from Step-2. Number of keywords should not be limited. However, psychologically, the capability of cognition is limited and necessary to focus on a limited number of keywords for use in the model. In our experience, model elements should be limited to fewer than 15. Otherwise, people feel model is too complicated and have difficulty understanding the whole structure.
We ask evaluators to prioritized keywords and send their answers via the Internet. Based on those answers, we select the top 10-15 as basic standardized keywords that would represent the public policy of

Tatebayashi City.

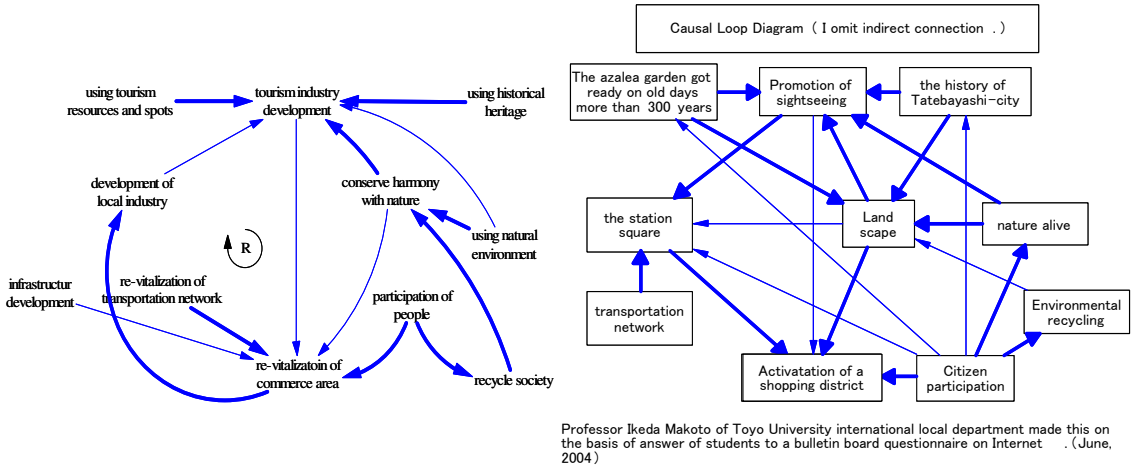


Figure 6: Sample ST model for evaluation (with Vensim and SumTaKN)

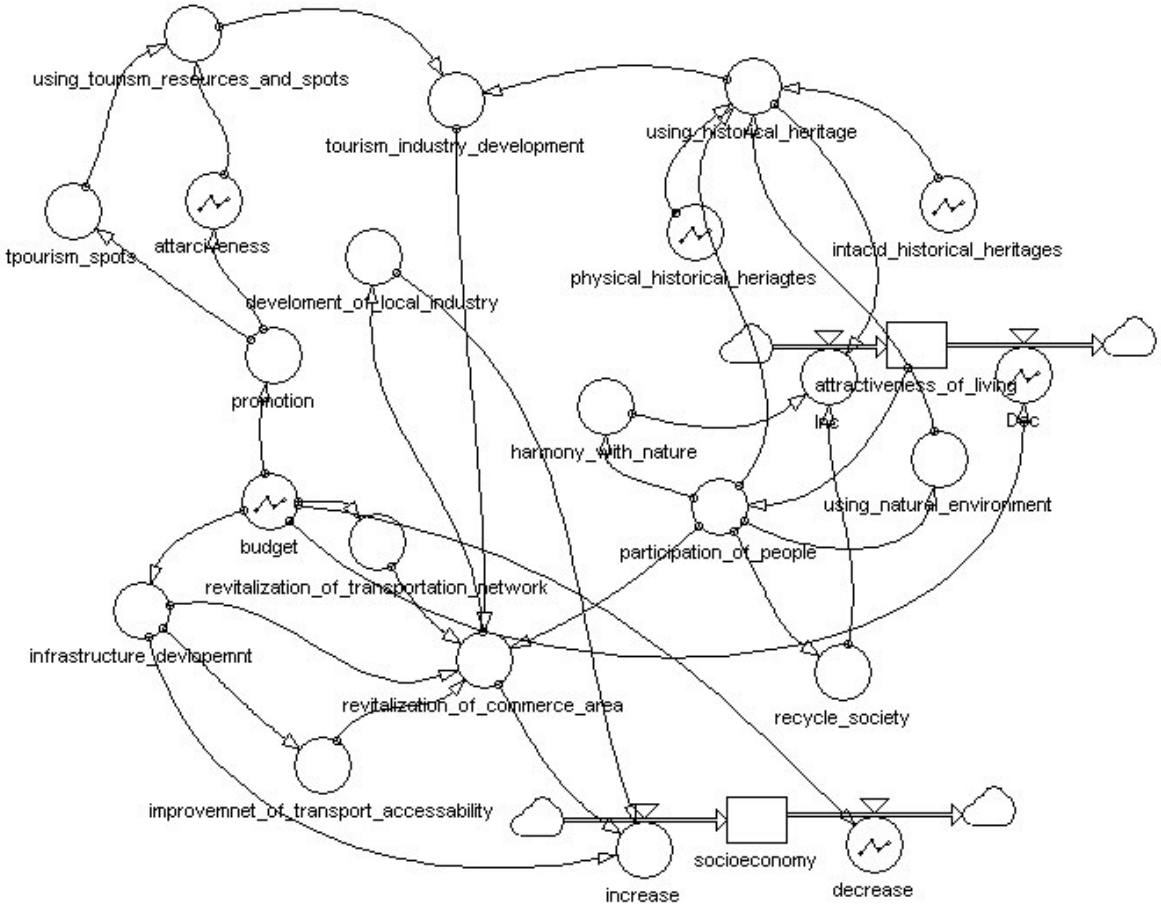
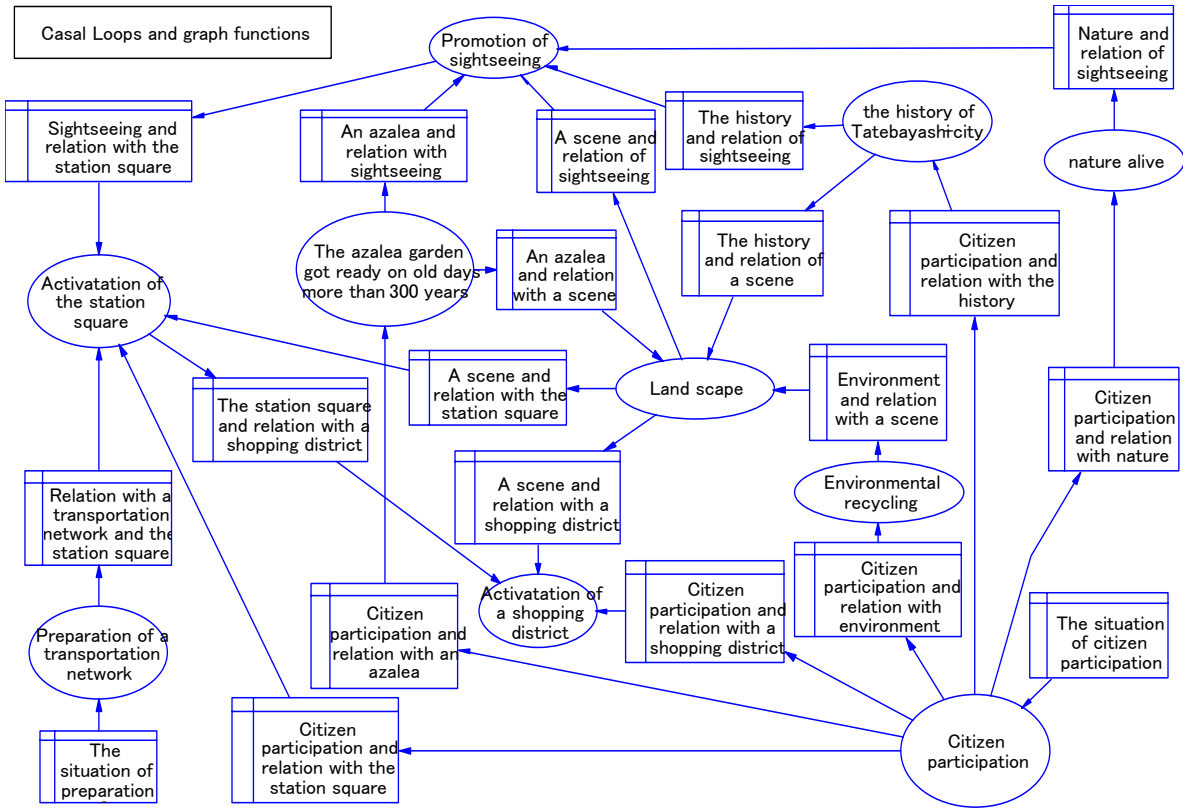


Figure-7a: SD model for evaluation (sample developed by Powersim)



Professor Ikeda Makoto of Toyo University international local department made this on the basis of answer of students to a bulletin board questionnaire on Internet(June, 2004)

Figure-7b: Dynamic ST model for evaluation (sample developed by SimTaKN)

Step-4: Development of ST model

We request evaluators to make relation between 10-15 selected basic keywords. Some elements (keywords) have multi relations and some have single relation. After completion of relations, then ask evaluators to draw behavior graph to explain relation of two elements. After that, the evaluator can make ST model for considering their ideal grand policy. This technique of determining relation of elements with behavior graph is widely accepted in business strategies such as Kim Warren (2002).

Figure-6 shows one model prepared by an evaluator to explain his idea for concentrated enforcement of re-vitalization of commercial area with encourage participation of citizens.

Step-5. ST simulation

We then request evaluator to make ST simulation based on the ST model built in step-4 in terms of what ideal grand policy should be. Also, request what behavior graph would be changed with their grand policy.

Step-6: More SD modeling

Evaluator feels they need to put more elements in their model to explain behavior of keyword more precisely. We encourage inserting more elements to their model to explain more accurate behavior of model. As a result, model becomes more complex and somehow difficult to explain, though, behavior of model is basically the same with ST model composed with 13 elements built in Step-4 and 5. If concept is rapidly changed, we suggest going back to Step-4 and 5 and change some keywords, if necessary. However, if concept is not changing rapidly, we also suggest keeping ST model and use only for explanation.

Figure-7 shows sample of one model that evaluator created where he inserts more elements and most relations are determined with table functions. This model is quite high level and he could not determine detail of policies and projects. However, it may not be necessary to know detail of projects and plans for purpose of evaluation simulation.

We do not limit modeling tools. One of evaluator has rich experience with Vensim and Powersim. He wishes to develop the model with Vensim and Powersim. However, other evaluators do not have much

experience with SD modeling. For those evaluators, we recommend to using SimTaKN (M. Ikeda, K. Nakamura and T. Suetake 2004) as tool for beginner. This tool makes ST dynamic model much easier.

Step-7: Evaluation

Lastly, we request doing simulation of their ideal grand policy in model built in Step-6. In this stage, the model is still between ST model and SD model. Many relations are determined as graph, or table. We encourage evaluator to determine relations more rigorously with mathematic equations. Any SD simulation software can describe this model, although the results may include too much table function and some SD practitioners do not like this in such a model. Evaluator can conduct simulation and obtain conclusion of grand policy. Again, we asked them to send the results and compared it with the Internet.

We evaluate adequacy or proprieties of present grand policy in term of it is best alternatives that evaluator can possibly consider. We collect simulation results and made table of several major performance indicators including development of tourism industry (represent with sales), and budget.

Figure 8 shows some sample of evaluation results. List up and compare with simulation results of evaluator's models, we can conclude adequacy or proprieties of public policy with diversity of view.

Using ST/SD model, we can solve the difficulty on issue of rationale as well as impact and effects. Evaluation practitioners can understand model composed with 10-15 keywords. Also, they can accept simulation results from more rigorous ST/SD model.

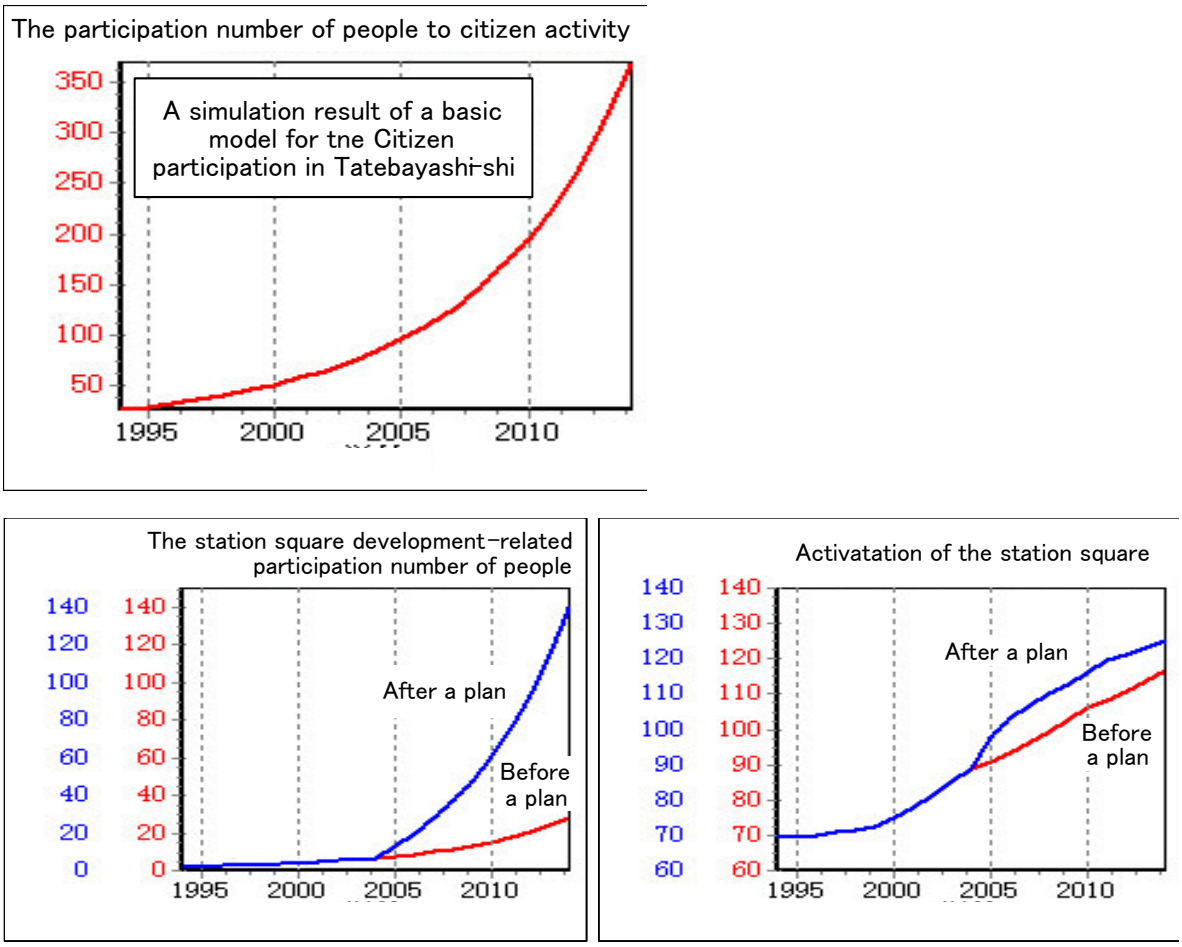


Figure 8: Sample results of evaluation with simulation

1) Objectives Achievement:
- Target goal: Citizen participate **400,000 to 500,000**
- Actual performance: **250,000**
2) Impact and Effects:
- simulation result of without policy: **0**
- simulation result of after policy: **350,000**
3) Efficiency:
- focus on citizen participation; citizen participation is **350,000**
- focus on direct tourism ind. dev.; **100,000**
- focus on direct commercial area revitalization: **150,000**
4) Rationale (relevance/adequacy):
- model focus on citizen participation; citizen participation is **350,000**
- model focus on direct tourism ind. dev.; participation is **little**
- model focus on direct commercial area revitalization: **little**
5) Sustainability: simulation result of long term period; citizen participation is **increasing**

Table 1: Comparison of simulation results with each evaluator's quantitative model

4. Further study

The result of grand policy of Tatebayashi municipal government is rationale with simulation results. Some model rather focus on impact of infrastructure and another focus on tourism industry development. Model builder develop variety models and this diversity gives evidence for evaluation of rationale. Also, best model shows another evaluation issues such as impacts and effectiveness.

We show the table of simulation results and several best models to Tatebayashi municipal government. Those models show weakness and missing links within six basic policies for supporting four basic concepts. They basically agree with the evaluation results and also SD/ST model itself gives significant impact of their evaluation process. They request to Professor Ikeda for developing SD/ST model for evaluation before conducting questionnaire survey and design public policy evaluation. Several best models suggest further research is necessary when they renewing and updating their grand public policy.

Further steps are needed to properly promote this technique to evaluation practitioners. In fact, policy evaluation by local governments was not previously considered regarding the 5 evaluation issues in the evaluation principles of DAC. The author believes it is necessary to consider the 5 issues in the evaluation process for policies of local governments. It is widely accepted for evaluation of government-to-government assistance projects and programs, which may be first step. However, we encourage conducting an evaluation using group modeling techniques mentioned previously in this paper.

5. Comments for Basic Concept

Evaluation is determined as, activity to get judgment or conclusion from results of measure on model that is possible to recognize as represent with target. Normally and specially in socio-economic development issues, target is very difficult to measure directly, or insufficient. Therefore, make a model that could be assumed as represent. It may be common sense for System Dynamist. Also, appropriateness of model as represent with target is agreement with model builder and model user. We do not argue with this concept of model and model user.

If evaluator may agree with a model represent with target of evaluation, then evaluator can measure result of simulation select all possible parameter combination. In this situation, model builder and model user focus on their interest to build best model. Many discussion focus on in this issue, how to build best model, and many methodology such as group modeling is developed.

However, there are also cases as evaluator could not agree with one model is enough for represent

with target. In this situation, some methodology may need to build all possible models that could cover all possible representation with target.

6. Bibliography

FASID: Foundation for Advanced Studies on International Development, PCM: Project Cycle Management – Management Tool for Development Assistance, FASID, 1999

FASID, Monitoring and Evaluation with PCM: Project Cycle Management Methodology, FASID, 2000

M. Ikeda, K. Nakamura and T. Suetake, New Tool for Improving Planning Capability of Local Government Staff, ISDC2004 Proceeding, 2004

M. Ikeda, Eco-Symbiosis Sociology, Asakura, Tokyo 2004

M. Ikeda, New Direction for Sustainable Regional Re-Vitalization in Environment and Tourism, Report of Symposium, Institutes of Regional Re-vitalization Study, Toyo University, 2004

M. Ikeda, Study on Regional Re-Vitalization with Improving Traffic System, Institutes of Regional Re-vitalization Study, Toyo University, 2004

Y. Itakura et al, Evaluation of AOTS Training Program, Japan Evaluation Study V2 N2, Japan Evaluation Society, 2002

G. Kelly and P. Walker, Community-Level Systems Thinking, ISDC2004 Proceeding, 2004

K. Kishi, et al. Study on Effectiveness of Water Supply System in Developing Countries to Social Issues – Process of Effective Appears, Proceeding of First National Conference, Japan Evaluation Society, 2004

Y. Morita et al. Next Step of Policy Evaluation with True Management Tools, Japan Evaluation Study V4 N1, Japan Evaluation Society, 2004

M. Pedercini, Evaluation of Alternative Development Strategy for Papua Indonesia: A Regional Application of T21, ISDC2004 Proceeding, 2004

Kim Warren, Competitive Strategy Dynamics, Wiley and Sons, July 2002

J. M. Watkins, B. J. Mohr, Appreciative Inquiry: Change at the Speed of Imagination, Jossey-Bass, 2001