# 日本評価研究

# Japanese Journal of Evaluation Studies

Vol. 11, No. 2, November 2011

#### 特集:地方自治体ベンチマーキング:組織間比較の 意義と可能性

特集:地方自治体ベンチマーキング:組織間比較の意義と可能性

田中 啓 小野 達也

自治体評価におけるベンチマーキングの可能性:

米国の取り組みが示唆するもの 田中 啓

自治体ベンチマーキングと指標の比較可能性 小野 達也

都市行政評価ネットワーク会議

自治体業務改善のためのベンチマーキング 熊倉 浩靖

川西市における行政評価の取り組みとベンチマーキングの可能性

本荘 重弘

福井市の行政評価と自治体ベンチマークシステム

中村 葉子

行政部門評価におけるベンチマーキングの現実機能

一評価主体の協働と評価性向試論ー

西出 順郎

#### 研究論文

行政評価過程における職場討議の導入とその効果検証

佐藤 徹

技術協力プロジェクトの効果に及ぼすガバナンスの影響

菊田 怜子 牟田 博光

### 日本評価学会 Japan Evaluation Society

#### 『日本評価研究』編集委員会 Editorial Board

編集委員長 三好 皓一(立命館アジア太平洋大学)

Editor-in-chief Koichi MIYOSHI

副委員長 西野 桂子(ジーエルエム・インスティチュート)

Vice-Editor-in-chief Keiko NISHINO

常任編集委員 牟田 博光(東京工業大学)

Standing Editors Hiromitsu MUTA

山谷 清志(同志社大学) Kiyoshi YAMAYA

編集委員 青山 温子(名古屋大学) 岩渕 公二(NPO政策21) Editors Atsuko AOYAMA Koji IWABUCHI

大島 巖(日本社会事業大学)

Iwao OSHIMA

岡本 義朗(三菱UFJリサーチ&コンサルティング(株))

Yoshiaki OKAMOTO

小野 達也(鳥取大学) 窪田 好男(京都府立大学)

Tatsuya ONO Yoshio KUBOTA

佐々木 亮(国際開発センター) 佐藤眞理子(筑波大学)

Ryo SASAKI Mariko SATO

渋谷 和久(国土交通省) 鈴木 絲子(神戸女子大学)

Kazuhisa SHIBUYA Itoko SUZUKI

田中 弥生(大学評価・学位授与機構) 西出 順郎(岩手県立大学)

Yayoi TANAKA Junro NISHIDE

林薫(文教大学)松岡 俊二(早稲田大学)Kaoru HAYASHIShunji MATSUOKA

源 由理子(明治大学) Yuriko MINAMOTO

事務局 〒140-0002 東京都品川区東品川4-12-6 日立ソリューションズタワーB 22F

Office 一般財団法人国際開発センター内

特定非営利活動法人日本評価学会

編集事務担当 鈴木 直美

TEL: 03-6718-5931, FAX: 03-6718-1651 E-mail: jes.info@evaluationjp.org

# 日本評価研究

### 第11巻 第2号 2011年11月

## 目 次

田	:地方自治体ベンチマーキング:組織間比較の意義と可能性 中 啓
	中 啓 自治体評価におけるベンチマーキングの可能性:米国の取り組みが示唆するもの3
	野 達也 自治体ベンチマーキングと指標の比較可能性13
	倉 浩靖 都市行政評価ネットワーク会議 自治体業務改善のためのベンチマーキング31
	荘 重弘 川西市における行政評価の取り組みとベンチマーキングの可能性45
-	村 葉子 福井市の行政評価と自治体ベンチマークシステム
	出 順郎 行政部門評価におけるベンチマーキングの現実機能 -評価主体の協働と評価性向試論- ·······················65
	<b>藤 徹</b>
3	行政評価過程における職場討議の導入とその効果検証77

菊田 怜子 年田 博光 技術協力プロジェクトの効果に及ぼすガバナンスの影響	91
春季第8回全国大会開催報告 春季第8回全国大会プログラム(実績)・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	110
<b>委員会活動</b>	115
広報委員会・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	119
日本評価研究刊行規定······日本評価研究投稿規定·····日本評価研究投稿規定····日本評価研究執筆要領····日本評価研究査読要領·····日本評価研究査読要領·····	122
Publication Policy of the Japanese Journal of Evaluation Studies	131

#### 【巻頭言】

特集:地方自治体ベンチマーキング:組織間比較の意義と可能性

特集にあたって

田中 啓 静岡文化芸術大学

小野 達也 <sup>鳥取大学</sup>

現在、地方自治体の過半数(都道府県と市に限れば8割以上)が行政評価制度を導入しており、しかも、これまでその数は一貫して増え続けてきた。こうした状況を踏まえると、少なくとも見かけ上は、自治体(特に広域自治体と都市自治体)にとって行政評価は「標準装備」となりつつあると言えよう。

その一方、行政評価の実施に関わる問題点が指摘されるようになってからも久しい。「評価指標を適切に設定することが難しい」「評価がやりっぱなしになっており、結果が利用されていない」「評価の作業負担が重く、職員が『評価疲れ』している」といった状況は、程度の差こそあれ、行政評価に取り組む多くの自治体が直面する現実である。

実務の現場や学会においては、こうした問題の解決方法について、ひと頃盛んに議論が行われたが、残念ながら、現実が大きく改善したようには見えない。むしろ近年は、関係者の間でも行政評価に関する話題が以前ほどは取り上げられなくなっており、「評価離れ」(自治体関係者の評価に対する関心の希薄化)が懸念される状況である。

本特集では、こうした行政評価の閉塞状況に1つの突破口を開く可能性のあるものとして「ベンチマーキング」に注目する。ベンチマーキングとは、元々は民間企業が経営改善や戦略立案に用いるために開発された手法であるが、近年では行政機関や非営利組織においても応用が進んでいる。

ベンチマーキングにおいては、自組織の業績を他組織(同業種の優良企業など)の業績と比較し、そこから新たな知見を得ることが主眼である。この「自組織を他組織と比較する」という観点は、従来の行政評価においてはあまり重視されてこなかったものである。だが、こうした「組織間比較」の観点を導入することにより、従来の行政評価に新たな意義が見出される可能性がある。

このような認識に基づき、本特集では、自治体がベンチマーキングを実施することの意義や可能性について、直接的または間接的に考察した論文を掲載した。

まず田中論文は、主にアメリカの地方自治体におけるベンチマーキングの取り組みを詳しく紹介するとともに、アメリカの取り組みが日本に対して何を示唆するのかを論じている。

小野論文は、ベンチマーキングの中核である業績指標に焦点を当て、自治体間・地域間で指標を比較するための条件を明らかにし、その観点から各種比較方法の課題と可能性を理論的に考察している。

熊倉論文は、国内の代表的な自治体ベンチマーキングの取り組みである「都市行政評価ネットワーク会議」について、その概要を紹介するとともに、具体的な指標を示し、計測結果から得られる含意を論じている。

本荘論文は、川西市における行政評価への取り組みの経緯と現状を総合計画との関連性も含めて詳しく紹介した上で、川西市におけるベンチマーキングの位置づけや運用上の課題を検討している。

中村論文は、福井市独自のベンチマーキングの取り組みである「比ベジョーズ」について、その導入の 経緯、位置づけ、実施方法を詳しく解説した上で、「比ベジョーズ」の成果や課題を考察している。

日本評価学会『日本評価研究』第11巻第2号、2011年、pp.1-2

最後に西出論文は、評価に関わる多様な評価主体に注目し、それぞれの評価主体の持つ傾向(評価性向)を理論的に導いた上で、ベンチマーキングにおいて不可欠となる評価主体間の協働の可能性について論じている。

以上で紹介したように、本特集には、ベンチマーキングの意義・手法・現状・課題等について理論面あるいは実務面から論じた論文がバランスよく配置されている。日本において、公共部門のベンチマーキングについての実務や研究はまだ緒についたばかりであるが、本特集が契機となり、この分野の実践及び研究が一層盛んになれば幸いである。

#### 【研究論文】

### 自治体評価におけるベンチマーキングの可能性: 米国の取り組みが示唆するもの

田中 啓 静岡文化芸術大学 hiraki@suac.ac.jp

#### 要約

行政評価の有効性を高めるための方策はいまだ試行錯誤の段階であるが、米国で開始された自治体間ベンチマーキングは、行政評価に新たな可能性を拓く手法として期待される。米国の代表的な自治体間ベンチマーキングであるICMAの比較業績測定プログラムには、160を超える自治体が参加しており、参加自治体は適切な自治体との業績の比較が可能となっている。自治体間ベンチマーキングは、新たな比較対象を導入することにより、業績測定をより有意義なものとしてくれるものであり、自治体間の同質性が高い日本においては、自治体間の比較を行うことは米国以上に意味を持ちうる。自治体間ベンチマーキングは既存の行政評価制度を基礎として取り組むことができるため、比較的小さいコストで実施することができるという利点を持つが、その取り組みを真に有効なものとするためには、多数の自治体の参加、事務局機能を果たす専門的組織の存在、安定的な資金源の獲得がポイントとなる。

#### キーワード

自治体評価、業績測定、ベンチマーキング、比較業績測定

#### 1. はじめに

国内の地方自治体が行政評価に本格的に取り組み始めてから十数年が経過した。この間に行政評価制度を導入し、実施する自治体の数は着実に増加している。その一方、行政評価制度を導入したにもかかわらず、制度の実施に際してさまざまな問題に直面し、行政評価を導入した効果を発揮させることができずにいる自治体は少なくない。

行政評価制度が抱える問題点に対しては、各地 の自治体がさまざまな対策を検討しているが、い まだに決定的な解決策が提示されているとは言い 難い。自治体の行政評価は、その普及度が向上したにもかかわらず、成熟の段階には程遠く、いまだに学習や改善が必要な取り組みであり続けている。

本稿では、こうした国内の現状に鑑み、自治体評価の有効性を高める可能性のある手法として、 米国の自治体が実施している「自治体間ベンチマーキング」に注目する。自治体間ベンチマーキング」に注目する。自治体間ベンチマーキングとは、複数の自治体が連合を組み、自治体間に共通の行政サービスや活動についての業績情報を共有し、連合に参加する自治体が個々の改善に利用することを目的とする手法または取り組みであ 4 田中 啓

る。

自治体ベンチマーキングは米国で1990年代半ばに始まった取り組みであり、最も代表的なものは、国際郡市経営協会(International City/County Management Association、以下ICMAと表記)による自治体間ベンチマーキングである。全米規模で展開されているICMAによる取り組みの他にも、特定の州や地域に限定して実施されている局地的な取り組みも存在する。こうした米国の取り組みに触発されて、日本でも近年、類似の取り組みが開始されている。

そこで本稿では、日本の自治体評価において、ベンチマーキングを実施することの意義や可能性 について考察を行う。

以下では、まず2において、日本の自治体における行政評価の普及状況と問題点について概観する。続く3においては、ベンチマーキングの基本的な考え方を整理した上で、米国の自治体におけるベンチマーキングの取り組みの状況を整理する。さらに4においては、自治体間ベンチマーキングのうち特にICMAによる取り組みに焦点を当てて、その実態と特徴を検討する。最後に5において、自治体ベンチマーキングは日本の自治体評価にとってどのような点が有意義であるのか、さらに日本で自治体間ベンチマーキングを実施する上でどのような課題があるのかという点について検討を行った上で、日本における自治体ベンチマーキングの可能性を論じて結語とする。

#### 2. 日本における自治体評価の現状

#### (1) 自治体評価の普及状況

1990年代後半以降、国内の自治体においては、行政改革の一環として行政評価の導入が進んでいる。特に都道府県と市・特別区における導入が進んでおり、総務省の調査結果(総務省 2011)によれば、2010年10月1日時点で都道府県の98%(46団体)、市区の80%(651団体)が行政評価制度を導入済みである。一方、町村においては、行政評価を導入している自治体の割合は30%(280団体)に留まっているものの、行政評価を導入する自治体の数は毎年着実に増加している。

行政評価制度を導入する自治体が増加した結果、行政評価を導入済みの自治体数は977団体、全自治体における行政評価の普及率は54%に達している。行政評価をまだ正式に導入していない自治体においても、既に行政評価制度を試行中のところや導入を検討中のところが多数存在することから、今後も行政評価の普及率が徐々に高まっていくことは、ほぼ確実な情勢である。

#### (2) 自治体評価の問題点

このように国内の自治体に行政評価が普及する 一方、行政評価に取り組んでいる自治体において は、行政評価の問題点が認識されるようになって いる。

#### ①実施の難しさ

多くの自治体が導入している行政評価においては、業績測定(Performance Measurement)が主に利用されている。業績測定とは、事業や施策別にその業績を測定するための評価指標を設定し、指標ごとの実績値と目標値を比較することにより評価を行う手法である。

しかし、これを適切に実施することは必ずしも容易なことではなく(田中 2008a)、多くの自治体は、評価指標の中でも特に事業や施策の成果を示す成果指標や指標ごとの目標値を設定する際に技術的な難しさに直面している。また、そもそも個々の事業・施策の目的や全体的な施策・事業の体系が適切に設定されていなければ、事業や施策を対象として評価を行っても有意義な結果を得ることが難しい。

#### ②作業負担・運用コストの大きさ

行政評価においては、評価シートなどと呼ばれる共通の様式に事業や施策ごとの評価結果を記入する方法が一般的である。多くの自治体が導入している事務事業評価の場合、評価対象となる事務事業数は数百から千以上に及ぶ場合もある。評価シートに記入するだけでも、多数の職員がこの作業に従事することとなり、庁内全体では膨大な量の作業が発生することになる。また評価シートに記入された内容を庁内で回覧あるいは対外的に公表するために、その内容を取りまとめる作業も発

生する。もちろん、本来の評価活動であるデータ の収集や分析作業も、知識や経験のあるなしにか かわらず、多大な時間を必要とする作業である。

#### ③評価の有効性に対する不信

このように、行政評価を導入している自治体においては、職員が膨大な時間と労力をかけて評価活動を行っている。しかし、職員が手間暇をかけて評価を実施しているにもかかわらず、その評価結果は十分に活用されているわけではない(田中2008a)。

評価結果が活用されない(あるいはこれを活用しようとしてもできない)場合、その自治体において行政評価の有効性や有用性に対する職員の不信感が生じることにつながる。評価結果が活用されていないとすれば、行政評価がその効果を発揮することはできないはずだからである。

行政評価制度の有効性に対する職員の信頼感が 揺らいだ場合、「そのような行政評価制度は、早 晩自治体内部で職員が真剣に取り組む対象ではな くな」り(田中 2008 a、p.53)、その評価制度の 有効性がさらに毀損されるという悪循環に陥るこ とになる。

#### (3) 自治体の対応

このような状況を打開するために、自治体の中には、従来の行政評価制度に大きな変更を加えているところがある。例えば、北海道、新潟県、三重県、横須賀市といった自治体は、いずれも先駆的に行政評価に取り組んできたところであるが、近年、従来の行政評価制度を大きく改訂している<sup>2</sup>。また既存の行政評価制度の基本構造を維持した上で、従来の事務事業に加えて施策レベルや政策レベルの評価へと評価対象を拡大することにより、行政評価制度の有効性を高めようとする自治体は数多く存在する。

残念ながら、自治体のこうした取り組みが功を 奏していると判断するだけの材料は今のところな い。今後もしばらくの間は、行政評価を導入した 自治体において、さまざまな試行錯誤が続けられ ていくものと考えられる。

こうした状況に対して、自治体間ベンチマーキングは新たな可能性を拓いてくれる手法であるい

うのが本稿の主張である。そこで以下では、自治 体間ベンチマーキングに先駆的に取り組んできた 米国の実態を検討する。

#### 3. 米国における自治体間ベンチマーキング

#### (1) ベンチマーキングとは

評価活動においては、注目する事象についてデータを収集し、そのデータを分析することにより何らかの判断を行う。判断を行うためには、収集したデータを照らし合わせる何らかの比較対象が必要である。

業績測定を例に取れば、施策の効果を示す評価 指標(これを成果指標またはアウトカム指標と呼 ぶ)を設定し、その実績値を計測することになる が、実績値を把握するだけでは判断を下すことが できない。実績値の水準が高いのか低いのか、望 ましいのか望ましくないのかを判断するための基 準が存在しないからである。このため何らかの比 較対象を設定することが必要となる。

比較対象を設定する方法として代表的なものは、次の3通りの考え方である。

- ①自己の過去の水準を比較対象とする
- ②先見的に与えられた到達目標を比較対象とする
- ③他の主体の水準を比較対象とする

①は評価対象自身の過去の水準を比較対象とする考え方であり、必然的に時系列的な比較作業になる。実際、評価を行う際に過去の水準と比較することは日常的に行われており、最も一般的な方法である。

②は、関心の事象についてあらかじめ設定されている目標水準を比較対象とする考え方であり、環境基準のように目標値が法令で定められている場合や、選挙公約のように政治的に目標が設定されている場合もある。②の手法も業績測定において広く採用されている。

最後の③は、①、②ほど一般的ではないが、比較対象として当然考慮されるべき方法である。これは、比較対象を他の主体に求めるものであり、

6 田中 啓

特定の施策の効果が評価対象の場合であれば、他の部局の類似する施策や他の自治体の施策を比較対象とする場合などがこれにあたる。時によっては、民間企業を比較対象とする場合すらある。このように自己以外の組織の業績等の水準に注目する手法がベンチマーキングである。

「ベンチマーク」とは、業績面で参考にすべき 基準を意味する。組織内外に「ベンチマーク」と なる比較対象を見つけ出し、その比較対象と自組 織の業績を比較することにより、自組織の問題や 劣っている点を明らかにする手法がベンチマーキ ングである。

ベンチマーキングにおいては、ベンチマークとなる比較対象のことを「ベスト・プラクティス」と呼ぶ場合もある。この場合は、関心の対象分野における「最善事例」と自己を比較することにより、当該分野で最高の水準を目指していくという意味合いが強くなる。

ベンチマーキングは元々米国の民間企業向けに 考案された手法であるが、近年は行政機関や非営 利組織においても利用されるようになっている。 この手法のメリットとしては、追いつくべき「他 者」を比較対象とすることにより、組織として業 績の改善に向けての明確な意識づけができること や、単なる業績の比較にとどまるのではなく、比 較対象の実践方法に学ぶことにより、自己の業績 を向上させる具体的な方策を検討できることを挙 げることができる。

#### (2) 行政部門におけるベンチマーキングの事例

米国において、ベンチマーキングは民間部門では広く普及しているものの、行政部門においてはまだ普及度が高いとは言えない。しかし、近年この分野で注目すべき取り組みが実施されている。

ベンチマーキングへの取り組みは大別すると2つのタイプがある。

1つは自治体間ベンチマーキングと呼ぶことができるものであり、複数の自治体³が連合を組んで、それぞれの業績に関するデータを共有し、加盟自治体間でベンチマーキングを行うタイプのものである。その代表としては、ICMAによる比較業績測定プログラム(Comparative Performance Measurement Program)やノース・カロライナ・ベ

ンチマーキング・プロジェクト (North Carolina Benchmarking Project) を挙げることができる<sup>4</sup>。

2つ目のタイプとしては、行政部門における優良事例の認定や表彰に重きを置く取り組みがある。これは、あらかじめ設定された項目により組織の全般的な特性や行政運営の質を評価するものであり、評価結果は得点化されるほか、優良な行政機関に対して表彰を行う場合もある。代表的な事例には、マルコム・ボルドリッジ賞(Malcolm Baldrige National Quality Award)やハーバード大学による米国行政機関改革賞(Innovations in American Government Awards)などがある。

このタイプの取り組みは、ベンチマーキングを 行うこと自体ではなく、ベンチマーキングの対象 となりうる優良事例の探索に重点が置かれている ため、厳密な意味ではベンチマーキングと呼ぶこ とはできないかもしれない。しかし、あらかじめ 公表された評価項目を用いて自組織と優良事例と を比較することが可能であるという点で、ベンチ マーキングに類似の取り組みととらえることがで きる。

これらの2つのタイプの取り組みのうち、自治体間ベンチマークは従来の日本の自治体における行政評価を拡張するものととらえることができる。一方、2番目のタイプは組織評価としての性格が強く、従来の行政評価とは別の文脈の取り組みと考えられる。本稿の冒頭で指摘したように、国内の自治体の多くが行政評価の活用面で問題を抱えているという現状に対して、自治体間ベンチマーキングは既存の行政評価の有効性を高める可能性を持っている。そこで、以下ではICMAの比較業績測定プログラムに焦点を当てて、その実態と特徴を検討する。

#### 4. ICMAの比較業績測定プログラム

#### (1) ICMAとは

ICMAとは、米国の市(city)や郡(county)の 幹部・管理職向けの専門組織である。設立は1914 年であり、現在の会員数は約9,000人である。

その活動領域は多岐にわたるが、地方自治体レベルの管理・運営に関するさまざまな分野に関す

る技術的支援や助言、教育・研修、情報・研究成果の提供等を行っている。「国際(International)」という言葉を冠していることから、国際的な視野を持つ組織であり、実際に外国人を会員として受け入れている。しかし、実質的な活動はアメリカ国内の自治体レベルに関わる内容が中心となっている。

ICMAでは自治体レベルにおける評価の実践を支援する取り組みも数多く行っているが、その中心となっているのがICMA内に置かれている業績測定センター(Center for Performance Measurement)という組織である。業績測定センターはさまざまな形で業績測定の普及啓蒙活動を行っている。中でも、近年特に力を入れている取り組みが比較業績測定プログラム(Comparative Performance Measurement Program)である。

#### (2) 比較業績測定プログラムとは

比較業績測定プログラムとは、前述のとおり自

治体間ベンチマーキングの取り組みであり、同種の取り組みの中では最も規模の大きなものである。

比較業績測定プログラムが開始されたのは1994年であり、当初の参加自治体(カウンティ、シティ、タウン、ヴィレッジ等)は44団体であった。現在の参加団体は163団体に拡大しており、規模別でみると人口規模が5万人未満の自治体が全体の5割強を占めている。ワシントンDCやカナダの自治体もこのプログラムに参加している。

比較業績測定プログラムでは、以下に挙げる18 分野を対象として、参加自治体が関連する業績等のデータを共有することになっている。ただし、全ての自治体が同じ行政サービスを提供しているわけではないことや、自治体ごとの固有の事情やニーズもあることから、参加する全ての自治体が18分野の全てにわたりデータを提供することを強制されているわけではない。。

表1 比較業績測定プログラムへの規模別の参加団体数

人口規模	参加団体数(%)	内 訳
10万人超	54 (33.1%)	市 36、郡 15、ワシントンD.C.、その他 2 <sup>造1</sup>
5~10万人	24 (14.7%)	市 19、郡 3、町 2
5万人未満	85 (52.1%)	市 67 <sup>22</sup> 、郡 1、町 11、村 5、その他 1 <sup>23</sup>
計	163 (100%)	

- (注1) その他の内訳は、Northwest Fire District (アリゾナ州)とSouth Florida Wataer Management District (フロリダ州)。
- (注2) カナダのセント・ジョンズ市(City of St. John's)を含む。
- (注3) Borough of State College (ペンシルヴァニア州)。
- (出所) ICMAのウェップサイト http://webapps.icma.org/cpmParticipants/?sort=locatio(2011年4月1日閲覧)より

#### 表2 比較業績測定プログラムの対象とする行政サービス分野

1. 規制の強制 10. 肥満防止 2. 建物・施設の管理 11. 公園とレクリエーション 12. 許可、土地利用、計画審査 3. 消防と救急医療 4. 車両管理 13. 警察 5. 道路の維持・補修 14. 調達 6. 低・中所得者向け住宅 15. 廃棄物とリサイクル 16. リスクマネジメント 7. 人材管理 8. 情報技術 17. 持続可能性 9. 図書館 18. 対若年層サービス

(出所) ICMAのウエップサイト http://icma.org/en/results/center\_for\_performance\_measurement/about/what\_we\_measure (2011年4月1日閲覧) より

8 田中 啓

#### (3) 比較業績測定プログラムの主要機能

ICMAの業績測定センターが比較業績測定プログラムの事務局機能を担っており、参加団体の募集・管理や参加自治体間の連絡・調整等を行っている。業績測定センターが事務局として提供する機能のうち特に重要なものは、次に挙げる5つの機能である。

#### ①データの収集・取りまとめ・報告

参加自治体の業績に関するデータを収集し、取りまとめることは、情報を共有するための最も基本的な機能である。データを収集するための共通様式 (テンプレート) は、分野ごとに用意されており、これらは参加自治体のメンバーによって作成された。

業績測定センターは、各参加自治体から共通様式に入力されたデータを受け取り、そのデータの内容をチェックする(これをデータ・クリーニングと呼んでいる)。チェックを受けたデータは、後述のようにウェッブサイト上に利用可能な形式で掲載されるほか、毎年、データを網羅的に掲載した年報としても発行されている。

一方、業績測定センターでは、収集したデータを加工・分析するサービスも行っている。例えば、自治体の個別の要望に応えて、全参加自治体における当該自治体の特定分野の業績の位置づけを示すグラフを作成して提供したりしている。

なお、比較業績測定プログラムで収集された自 治体別のデータを用いて、行政サービス分野別に 参加自治体のランキングを作成したり格付けを行 ったりすることは技術的には可能であるが、比較 業績測定プログラムとしてはランキングや格付け を一切行わないとしている。

#### ②ウェッブサイト上のサービス提供

比較業績測定プログラムの特徴として、ウェッブサイトを利用したデータの提供機能が充実していることを挙げることができる。業績測定センターのウェッブサイト上には、参加自治体のみがアクセスできる専用サイトへのリンクが用意されており、そこには基本的に全てのデータがダウンロード可能な形式で掲載されている。単にデータをダウンロードできるだけではなく、利用者はデー

タを検索したり、特定の自治体を複数選択して、 それらの自治体と自己の団体の業績に関するデー タを図や表で表示したりすることもできる。

これらの共有データに関する情報以外に、専用サイト上には、業績測定やベンチマーキングに関する解説や実践方法など、さまざまな情報や資料が用意されている。

#### ③技術的支援の提供

業績測定センターでは、業績測定やベンチマーキングの理論や実践方法について技術的な支援を必要とする自治体に対しては、直接担当者が自治体に出向いて研修活動などを行っている。このほか、業績測定センターは比較業績測定プログラムの運用に関わる全般的な技術的サポートを参加自治体に対して提供している。

#### ④コンソーシアム (Consortium) の結成支援

比較業績測定プログラムへの参加自治体のうち同じ地域に属する自治体や共通の関心を持つ自治体が参加し、参加自治体の目的や関心に照らして自治体間ベンチマーキングを独自に実施する取り組みがコンソーシアムである。コンソーシアムの開始を希望する自治体に対しては、業績測定センターが人的・技術的支援を提供している。現在、17のコンソーシアムが活動中である。

#### ⑤参加自治体間の交流

比較業績測定プログラムへの参加自治体のメンバーが参加する会議やワークショップが随時開催されており、参加自治体から業績測定やベンチマーキング等の実践例が報告される。また、こうした会議は、参加自治体間で情報・アイディアの交換を行う貴重な機会を提供している。

#### (4) 比較業績測定プログラムの特徴

以上の整理を元にしてICMAの比較業績測定プログラムの特徴を整理すると次のとおりである。

#### ①参加自治体が多いことによるメリット

比較業績測定プログラムにおいては、参加自治 体が160団体を超えている。類似の取り組みにお いて、これ程多くの参加自治体を集めている事例 はない。

このことは、自治体にとっては比較業績測定プログラムに参加する大きなメリットを提供している。つまり、参加自治体数が多いことにより、比較対象として適切な(例えば、自治体の規模や特徴が似ている、等の)自治体を探しやすい。また、参加自治体の平均値と自団体の実績値を比較する場合でも、参加自治体が少ない場合に比べれば、偏りの少ないより平均的な全体像と自組織とを比較することが可能となる。

#### ②ICMA業績測定センターによるサポート

比較業績測定プログラムには自治体が自主的に 参加し、その運営も参加自治体が主体的に行うこ とが期待されている。とはいうものの、比較業績 測定プログラムを円滑かつ有効に運用していく上 で、事務局である業績測定センターが果たしてい る役割は極めて重要である。参加自治体が米国全 土に分散していることから、参加自治体の結節点 の役割を果たす組織が必要であることは言うまで もない。

さらに重要なのは、業績測定センターが提供する技術的なサポートである。自治体間ベンチマーキングにおいては、参加自治体が共有するデータの質(定義の統一性やデータの信頼性など)が担保されなければ、ベンチマーキング自体の意義を見出しにくい。業績測定センターがデータのチェックを行ったり、参加自治体に対して技術的な支援を行ったりしていることは、比較業績測定プログラム全体のプロジェクトとしての信頼性や有用性を維持する上で重要な意義を持っていると言える。

# 5. 日本における自治体間ベンチマーキングの可能性

#### (1) 米国の取り組みの示唆

ICMAによる比較業績測定プログラムの現状を踏まえると、自治体間ベンチマーキングは以下の点で、日本の自治体の行政評価にとって有意義な取り組みであると考えられる。

第一に、自治体間ベンチマーキングは、評価に 新たな比較対象を導入することにより、業績測定 をより有意義なものにしてくれるというメリットがある。なぜならば、評価の比較対象を自己の過去の実績値や先見的に与えられた目標値だけでなく、外部の自治体に求めることができるからである。特に日本の場合は、米国に比べると、自治体の規模、組織体制、権限、提供する行政サービス等の面における同質性が高いことから、自治体間での比較を行うことが米国以上に意味を持ちうる。

また自治体間ベンチマーキングには、単に自治体同士の業績を比較するだけでなく、自己に比べて他の自治体の業績が良好である場合には、その自治体を仔細に観察することにより、自己の業績を向上させるためのヒントを得ることができるという利点もある。これは、取り組みが単独の自治体内部で完結する従来の行政評価では期待することのできないメリットである。

第二に、自治体間ベンチマーキングにおいては、 参加自治体がデータを持ち寄って比較しあうこと が核心ではあるが、参加者間のネットワークやコ ミュニケーションから得られる副次的なメリット を期待することもできる。例えば、データを比較 するだけではわからなかった他自治体との実践面 の違いが、その職員と直接コミュニケーションを 取ることにより把握できる可能性がある。

一方、ICMAの取り組みを参考にすれば、自治体間ベンチマーキングを実施する上では、以下の2点に留意が必要である。

まず、自治体間ベンチマーキングの有用性は、 参加する自治体が多いほど高まるという点である。これは、参加する自治体が多いほど、適切な 比較対象を見つけやすくなるという意味で、自治 体間ベンチマーキングに参加することのメリット が拡大するからである。言い換えれば、自治体間 ベンチマーキングを有意義なものとするために は、できるだけ数多くの自治体の参加を得ること が課題となる。

また、自治体間ベンチマーキングを円滑かつ有効に運用するためには、ICMAの業績測定センターが果たしたような、専門機関による不断のサポートが不可欠である。これは、各地に分散する自治体の結節点としての機能が必要とされるほかに、ベンチマーキングの有効性を担保するためには、データの質の確保や参加自治体に対する技術

10 田中 啓

的支援が重要だからである。

このような専門機関には、業績測定や統計解析の専門性を有する人材が必要であるし、このような人材を要して、参加自治体に対する支援や必要な活動を実施するためには、人件費を賄うための資金源の確保が課題となる。

#### (2) 日本における自治体間ベンチマーキングの 取り組み

これまで検討したように、自治体間ベンチマーキングを実施するためには、できるだけ多数の参加自治体を集めることや自治体間ベンチマーキングの活動を支援する専門機関の存在が不可欠である。こうした課題を解決することができれば、自治体間ベンチマーキングが日本の自治体評価にとって有意義な取り組みとなる可能性が高い。事実、国内には、その萌芽ともいえる自治体間ベンチマーキングの取り組みが既に存在している。

ICMAの比較業績測定プログラムに最も近い方法で実施されているのは、「都市行政評価ネットワーク会議」による自治体間ベンチマーキングである。この取り組みは、当初は総合研究開発機構(NIRA)のプロジェクトとして開始されたものであり、先立つ数年間の研究段階を経て2005年からベンチマーキングの実施段階に移行した。当初の参加自治体は15団体であったが、現在の参加団体は80団体を超えている。

参加自治体は20の施策領域の指標を共有しており、データの収集・分析や個別自治体に対する分析結果の提供には専門家が当たっている(熊倉2005)。ICMAの取り組みをモデルとしているだけあって比較業績測定プログラムとの類似性が強く見られるが、指標の選定やデータの分析方法は、日本の実態に合わせて洗練が図られている。

一方、都市行政評価ネットワーク会議に先立つ 2003年には、福井市を中心とする自治体間ベンチマーキングが開始されている。「比ベジョーズ」と呼ばれる取り組みで、福井市をはじめとする特例市34団体が参加している。比ベジョーズにおいては、17分野にわたる155指標が共有されており、インターネット上で利用可能なウェブシステムにデータが保管されており、参加自治体はこれを自由に利用することができる。また研究会や分科会

が定期的に開催されており、参加自治体間の意見 交換や交流の機会も提供されている。

なお、比べジョーズにおいては、ICMA業績測定センターのような専門機関は存在しておらず、参加自治体は運営経費を負担する必要がない代わりに、データ入力等の作業を分担しあうことが前提となっている。ただし、福井市が呼びかけて開始した取り組みであることから、システムの運営経費等は福井市が負担している。

#### (3) 日本における可能性

以上の整理を踏まえ、最後に日本における自治 体間ベンチマーキングの可能性について指摘したい。

まず、自治体間ベンチマーキングは、既存の行政評価に新たな比較対象を与えるものであることから、数多くの課題を抱える日本の行政評価にとって、取り組む価値のあるものだと考えられる。特に、多くの自治体が既に導入している事務事業評価をある程度基礎として取り組むことが可能であることから、現行の事務事業評価を維持しながら、比較的小さいコストで導入することができるというメリットがある。既に国内において都市行政評価ネットワーク会議や比べジョーズといった取り組みに数十団体の自治体が参加していることは、自治体間ベンチマーキングの有用性を示唆する事実と言える。

ただし、より有効な自治体間ベンチマーキングを行うためには、少しでも多くの自治体が参加することが望ましい。仮に多くの自治体の参加を集められたとすれば、ベンチマーキングの初期段階には、共有するデータの定義を決定したり、参加団体が使用する共通様式を作成したりする等のプロジェクト開始の準備のために、関係する主体がかなり大きな労力を投入する必要があるだろう。

このような導入初期の条件整備や運用開始後のプロジェクト管理や技術的サポートを十分に行うためには、ICMA業績測定センターのような事務局機能を果たす専門的組織の存在が不可欠である。残念ながら、日本にはまだICMAに相当する組織が存在しない。日本でICMAと同様の機能を担う可能性があるとすれば、シンクタンク、大学、学会、NPOといった機関であろう。

また、専門的組織のサポートを得て自治体間べ

ンチマーキングを継続的に実施していくためには、取り組みの経費を賄うための安定的な資金源が必要となる。自治体間ベンチマーキングに参加してメリットがあるのは参加自治体自身であることから、自治体が会費を分担しあうことが基本とされるべきである。筆者の計算によれば、仮に参加自治体が100団体とした場合、1団体が年間30万円ずつ負担することができれば、自治体間ベンチマーキングを運営していくために最低限必要な活動資金を確保することができる。

多くの自治体が財政面の制約を抱えている中で、新たな歳出を正当化することは決して容易なことではない。だが、既存の行政評価の実施においてこれまで浪費してきた行政資源の価値を客観的に算出してみれば、この程度の経費は決して大きいとは言えないことが明らかになるはずである。

#### 謝辞

本稿の完成にあたり、査読者の方から有益なご 指摘をいただきました。心から感謝いたします。

#### 注記

- 1 仮に総務省の調査(総務省2011)で「試行中」または「導入を検討中」と回答した自治体が全て行政評価制度を導入すると仮定すると、行政評価の普及率は93.4%まで上昇することになる。
- 2 行政評価制度の改訂方法を大まかに分類すれば、行政評価制度の実施目的を改める(例は新潟県)、採用する評価手法を当初のものから変更する(例は北海道)、評価対象を絞り込むなど行政評価制度の運用を従来より簡素化する(例は三重県、横須賀市)等がある。
- 3 米国について「自治体」という場合、シティ(市) やタウン(町)等のMunicipalityを指すが、これ以 後は説明の便宜上、「自治体」という言葉を本来の 自治体(Municipality)に加えてカウンティ(郡) も含めた意味で用いることとする。
- 4 これらの他にサウス・カロライナ州におけるサウス・カロライナ・ベンチマーキング・プロジェクト (South Carolina Benchmarking Project) やカンザス州におけるカンザス市地域業績測定プロジェクト

- (Kansas City Regional Performance Measurement Project) などの取り組みがある。
- 5 比較業績測定プログラムへの参加費用は、基本サービスを全て含む場合年間5,550ドル(1ドル80円で計算した場合、約44万円)である。ただし、人口規模が1万人未満の自治体や、行政サービス分野を限定してプログラムに参加する自治体の場合、別料金が適用される。
- 6 コンソーシアムの多くは同一州内の自治体によって 構成されている。例えばヴァージニア・コンソーシ アムには、チェスター・フィールド郡、フェアファ ックス郡、ジェームズ・シティ郡、プリンス・ウィ リアムズ郡、ダンヴィル市、ハンプトン市、ホープ ウェル市、ノーフォーク市、ポーツマス市(いずれ も比較業績測定プログラムの参加自治体)の7団体 が加盟している。

#### 参考文献

- 熊倉浩靖(2005)「都市行政評価ネットワーク会議 (仮称)に提供するテンプレート案」、『月刊NIRA政 策研究』、18(7):8-20
- 澤井安勇(2005)「都市行政評価ネットワーク会議 (仮称)設立の意義・目的」、『月刊NIRA政策研究』、 18(7): 44-47
- 総務省(2011)『地方公共団体における行政評価の取組状況』
- 田中啓 (2005)「米国における自治体経営とベンチマーキング」、『月刊NIRA政策研究』、18(7): 31-43
- 田中啓(2008a)「都市自治体の評価:本格普及から10 年後の実態」、『日本評価研究』、8(1):39-57
- 田中啓(2008b)「自治体評価の実像」、『公務改革の突破口』、東洋経済新報社
- 藤澤武明 (2005)「都市行政評価ネットワーク会議 (仮称)の設立手順・活動内容」、『月刊NIRA政策研 究』、18(7): 48-55
- 本荘重弘(2005)「行政評価システムを活用した行政 改善と今後の課題」、『月刊NIRA政策研究』、18(7): 21-30
- Ammons, N. D. (2001). *Municipal Benchmarks. 2nd Edition*. Thousand Oaks, CA: Sage Publications, Inc.
- Hatry, P. H. (1999). *Performance measurement: getting results*. Washington, D.C.: The Urban Institute Press.

12 田中 啓

Hatry, P. H. (2006). Performance measurement: getting results. 2nd Edition. Washington, D.C.: The Urban Institute Press

Keehley, P., Medlin, S., MacBride, S., & Longmire, L. (1997). Benchmarking for Best Practices in The Public

Sector. San Francisco, CA: Jossey-Bass.Morley, E., Bryant, S. P., & Hatry, P. H. (2001).Comparative Performance measurement. Washington,D.C: The Urban Institute Press.

(2011.8.26受理)

#### Future Prospects of Benchmarking in Local Government Evaluation: The Implications of Practice in the United States

Hiraki Tanaka

Shizuoka University of Art and Culture hiraki@suac.ac.jp

#### **Abstract**

Practitioners and researchers have been pursuing ways to improve the effectiveness of the evaluation practice of local governments in Japan, but such efforts remain unfulfilled. Comparative performance measurement, or benchmarking, which was developed in the United States, is regarded as a prospective method to provide local government evaluation with new possibilities. Over 160 municipalities participate in the ICMA's Comparative Performance Measurement Program, which is the most prominent benchmarking project in the United States. The participating municipalities benefit from the large membership, making it possible for them to compare their performance with that of the most relevant participating municipalities. Benchmarking, by enabling municipalities to introduce an additional object of comparison, makes current performance measurement practice more meaningful. This is especially true in Japan where there is less diversity among municipalities than in the United States. As a result, benchmarking among municipalities in Japan may be more suitable. Moreover, benchmarking is done with relatively small cost because municipalities that have already introduced performance measurement systems do not have to change or modify their systems when they start benchmarking. To make benchmarking efforts effective enough, however, participation of many municipalities, the intervention of an expert organization and a firm foundation of funds is necessary.

#### **Keywords**

local government evaluation, performance measurement, benchmarking, comparative performance measurement

#### 【研究論文】

### 自治体ベンチマーキングと指標の比較可能性

小野 達也 鳥取大学地域学部 onotatsu@rstu.jp

#### 要約

業績指標を自治体間で比較する自治体ベンチマーキングは、今後取り組みの拡大が期待される。自治体ベンチマーキングは、地域間の比較を行う社会指標との共通点も多いが、そこでの比較の実効性の鍵を握るのは、指標の比較可能性である。自治体ベンチマーキングの枠組みにおいて業績指標を自治体間・指標間で比べるために用いられる様々な方法について、指標値の尺度水準に着目しつつ、指標の比較可能性がどの程度得られるのかという観点からその課題と可能性を検討する。

#### キーワード

比較業績測定 (CPM)、比較可能性、尺度水準、標準化、調整済み業績指標 (APMs)

#### はじめに

行政評価・政策評価においては、どのような枠組みの評価であれ、政策・施策・事務事業の効果等を定量的に把握することが重視され、実際に今日では数多くの数・量の指標が用いられている。各所の評価作業においては、これら数・量の指標を用いた比較一時系列方向の比較、目標値と実績値の比較などが広く行われるが、今後の日本の公共政策・地域政策の検討・分析として、地域間の比較、とりわけ自治体間の比較一自治体ベンチマーキングがこれまで以上に重要になると考えられる。

数字の比較は、指標が比較可能であることが前提となる。地域間・自治体間比較においては、自治体の規模が様々であること一つをとっても、比較を意味のあるものとするために必要な手続きが

あることは明らかである。

一方、数・量の指標に基づく自治体間比較は、 社会指標の取り組みとして官民の双方に古くから の蓄積がある。そして、これらは自治体内の行政 評価・政策評価の拡張・応用としての自治体ベン チマーキングと、指標の比較作業において共通の 面がある。

ところが、地域間・自治体間比較において指標にはどのような比較可能性が求められるのか、社会指標の分野においても、必ずしもすべてが明確になっているわけではない。本稿はこのような観点から、まず内外における自治体間比較の現状を概観し、続いて指標の比較可能性という概念についてあらためて整理する。その上で、指標を比較するために用いられている様々な方法について検討を加え、その課題と可能性を論じたい。

#### 1. 指標による自治体比較の現状

近年、行政評価・政策評価の取り組みが大半の自治体に普及したが、その多くは所謂業績測定型の評価制度・システムである。そこでは政策・施策・事務事業のロジック(インプット~プロセス~アウトプット~アウトカム)の各段階をモニタリングする業績指標が設定され、定期的に計測されることとなる。

さて、業績指標をはじめする「指標」の多くは 比べることでその意味が明確になるものがほとん どである」。自治体の評価においては時系列の比 較と、そのヴァリエーションでもある目標管理即 ち目標値との比較(過去に将来の目標として設定 した値と現状・実績値との比較)²が専ら中心的な 役割を果たしているのが現状である。一方、時系 列比較と並ぶ統計情報比較の基本的方向であるク ロスセクション比較、すなわち業績指標の自治体 間比較³は余り行われていないが、近年自治体の 業績指標を比べる注目すべき取り組みが登場して いる。各自治体内における業績の測定とその活用 に加えて、他団体・地域とプログラム(政策・施 策・事業) の業績を比較することのメリットは、 他所との比較が競争をもたらす、ベストプラクテ ィスに学ぶ、また住民に業績比較情報を提供する (足による投票をもたらすこともあろう) ことな どを通じてマネジメントの向上・改善を促すこと であろう。

もっとも、行政評価・政策評価の業績指標として重視されるアウトカム指標は、いわゆる社会指標\*の指標と考えることができるものが多い。社会指標の地域(自治体間)間比較となれば、日本においても、官・民の両方において長年にわたる多種多様な事例がある。

アーバン・インスティテュートのハトリーらは、測定した業績情報を団体・地域間で比較する取り組みcomparative performance measurement(比較業績測定、本稿の以下ではCPMと呼ぶ)を、「消費者」指向の比較consumer-oriented comparisons、「活動団体」による比較advocacy group comparisons、義務付けられた報告・比較mandated reporting and comparisons、自主的共同による比較self-initiated cooperative comparisonsと

いう4通りに分類している(Morley et al. 2001)。 前三者は、業績を比較される団体が取り組みの推進主体ではないという意味で外部型externalであり、最後のものだけが、団体の積極的参加を伴うものである。ここでは、これら4種類のCPMをタイプ  $I \sim IV$  と呼ぼう。

タイプIの消費者指向の比較とは、民間の書籍 や雑誌が公共サービスの消費者すなわち住民に、 地域の状態や生活環境に関するデータを地域間比 較が可能な形で提供するものである。米国には、 北米のmetropolitan statistical area (MSA)単位で生 活の質Ouality of Lifeというという観点から、1981 年以降数年間隔で指標を作成・総合化してランキ ングをつけている "Places Rated Almanac"、 Money誌の "Best Places to Live"、Forbs誌の "America's Most Livable Cities"、Reader's Digest誌 の各種地域比較などがあり、日本でもそのランキ ングなどがしばしば報道されている。このような 「地域の暮らしやすさ」を測る企てとしては、世 界各都市の比較をする英Economist誌の"World's Most Livable Cities" や "Mercer's Quality of Living Survey"も広く知られている。また米国には、教 育や健康分野に特化した地域比較も多い(Morley et al. 2001)

日本でも民間版社会指標の多くがこのジャンルに相当するといってよい。週刊ダイヤモンドの全国都市ランキング、東洋経済新報社の全都市住みよさランキング(「都市データパック」に掲載)、朝日新聞社の都道府県別・都市別の民力指数などを継続的に刊行されているものの例として挙げることができる。また、旧経済企画庁が1992~1998年に、8生活領域別に都道府県別の総合指標を作成した「新国民生活指標」(People's Life Indicators, PLI)は、政府の作成する一種の社会指標でありながら、指標の作成意図・形態ともにこのジャンルに近いものといえる'。このように、都道府県や市町村を計測単位とする指標群から総合指標を算出する社会指標は、しばしば「地域総合指標」とも呼ばれている(大友 1997)。

タイプⅡの活動団体による比較とは、各種の活動グループ(またはロビイ団体、利害関係団体)などが、特定の社会問題等をテーマにCPMを作成・公表するものである。米国の例としては、

Governing誌の "Grading the States: A Management Report Card" などがある。日本では日経グローカル誌(旧日経地域情報)による全国市区の行政革新度ランキング・行政サービス度ランキング、週刊ダイヤモンド誌の全国都市の行革度ランキング、地方自治職員研修誌の「自治体の成績表」などが相当する。

タイプⅢの義務付けられた報告・比較とは、中央政府が地方政府に対し、あるいは(米国であれば)州政府が州内のコミュニティに対し業績情報の報告を求め、それらを比較する形で取りまとめるものである。英国のAudit Commissionにおいて1992年に開発され1993年に導入された、全地方政府が共通に計測・報告する業績指標が有名である。米国では、州レベルで特定分野における様々な取り組みがある(Morley et al. 2001)。日本では筆者の知る限り、該当する取り組みはない。

タイプIVのCPMが自主的共同による比較である。米国における継続的な取り組みの代表例として、International City/County Management Association(ICMA)の主導により1994年から取り組まれ、人口20万人以上の都市多数が参加するnational performance comparison projectとノースカロライナ大学のIOG(Institute of Government)が主導して1995年に開始されたCPM projectが挙げられる。。

日本において注目すべき事例が、都市行政評価ネットワーク会議である<sup>7</sup>。2001年度からの総合研究開発機構における研究を踏まえ2005年に15市が参加して発足、現在では参加市区町が80を超えている。また福井市が2002年に他の特例市に参画を呼びかけ、翌年度から運用されている「比べジョーズ」(現在は中核市に移行した団体も含め34市が協力)という興味深い取り組みもある<sup>8,9</sup>。

このほか、上記の4分類とは別に、各自治体が独自に他地域との比較を行う場合がある。例えば岩手県では2001年度より総合計画の体系に沿って都道府県別、県内の広域生活圏・市町村別の社会指標集(例えば岩手県2001a、2001bなど)を作成し、指標重視の政策評価システムの運用において積極的に活用していた。指標の比較方法を論ずる本稿の趣旨からはCPMの一形態と考えられる。

このように、業績指標を団体間・地域間で比較

するCPMには、目的・取り組み主体・比較形態 (指標を総合化するか否か、特定の観点・分野に 絞るか否か、など)において様々なものがあるこ とがわかる。日本において近年登場し、自治体内 の行政評価・政策評価として実行される業績測定 の拡張・応用と位置づけられる自治体ベンチマー キングは、これらCPMの部分集合であるといえ、 指標を地域間で比較するという場面においては、 いわゆる地域総合指標を含む他のタイプのCPM とも多くの論点を共有することとなる。

#### 2. 指標を比較するための条件一比較可能 性

#### (1) 妥当性、信頼性、比較可能性

CPMの成否は、当然のことながら、意味のある比較を行うか否か次第である。すなわち、指標は数字で表現されており、形式的に同じ形をしていれば最も素朴なレベルでの比較はいつでも誰にでも可能であるが、それが意味のある比較であるか否かは別問題である。素朴な比較は、間違った情報をもたらすという意味でしばしば有害ですらある。意味のある比較ができるか、しているかという条件は、一般に比較可能性comparability<sup>10,11</sup>と呼ばれている。

そもそも各種指標の取り扱いにおいては、妥当性validityと信頼性reliabilityが基本的要件とされる"が、Kalimo(2005)がOECDの社会指標における国際比較を論ずるにあたって適切に整理している通り、CPMにおいてはこの2条件に比較可能性を加えた3条件を(可能な限り)満たすべきであるといえる。

CPMにおける比較可能性の構成概念を検討する前に、指標の妥当性と信頼性について、自治体の業績測定の拡張・応用としてのベンチマーキング即ち各自治体が自らの指標を測定・算出してCPMを行う場合に、個々の自治体における業績指標の妥当性・信頼性とは異なる要請がCPMの比較可能性の観点からなされることを確認しておこう。

心理学や教育学の分野でなされてきた、測定・ 評価の原理としての妥当性の種類に関する議論を

参考にすれば、業績測定の指標に求められる妥当性は、プログラムのロジックのどの局面を測定するかという概念規定と指標設定に係る第1段階の妥当性(論理的妥当性)と、具体的に当該局面の何を測定するかという算式決定に係る第2段階の妥当性(操作的妥当性)とに分けて考えることができる<sup>13</sup>。

第1段階の妥当性について、自治体ベンチマーキングは当然どの指標についても、比較可能になるよう共通の概念規定を要求する。この点において、各自治体は自らの指標設定に至る事情と自治体ベンチマーキングの要求との間で折り合いをつけなければならない。すなわち、自治体における業績測定の制度・システムと自治体ベンチマーキングの関係をどのように構築・運用するかが課題となる。

第2段階の妥当性については、概念が統一された後に、いかに比較可能な操作的定義、算式を設定するか、という課題がある。個々の団体が測定・活用する業績指標そのままの形で自治体ベンチマーキングに使えればよいが、一般論として、比較を行う段階で何らかの加工等が必要になる場合も多いであろう。その場合やはり、各自治体における比較結果の解釈と運用に影響が及ぶこともあろう。

指標の信頼性については、信頼性の高い測定が望ましいという点で何ら自治体内の業績測定と変わるところはないが、信頼性の低いデータを提出する団体が1つでもあれば比較全体の有効性が損なわれる可能性がある。特に全団体のデータを用いて計算する相対評価、相対比較の指標を計算する場合、データの欠損や低い信頼性のデータは他団体の指標値に少なからぬ影響を直接与えることもあり得る。

さて、個々の団体が測定・算出して提出する自 治体ベンチマーキングの場合も、第三者が指標を 収集するCPMも、比較の対象となる指標が定ま ってからが、あらためて比較可能性が問われる場 面である。この比較可能性の内容を検討し、その 確保のための様々な手法を吟味することが、本稿 後半の目的である。

#### (2) 自治体ベンチマーキングに必要な比較可能 性

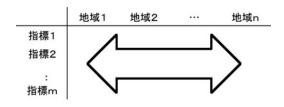
CPMで行われるクロスセクション比較は、図1に示すような(a)地域1,2,・・・,nの間で比較する、(b)指標1,2,・・・,mの間で比較する、という2通りの比較の組合せである。CPMの様々な取り組みは、これら2種類のクロスセクション比較の組合せ方により、次の2種類に分類することができる。

1つ目は、第1段階で個々の指標について自治体(地域)間の比較を行い、第2段階の比較として、個々の自治体(地域)において、地域間比較結果を指標間で比較する(指標1は相対的に○○の水準にあるが、指標2は相対的に△△の水準にある、など)ものである。これが「自治体ベンチマーキング型CPM」で、本稿1の分類のタイプⅢとⅣが相当する。

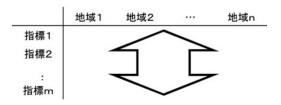
もう1つは、第1段階で個々の指標について地域間の比較を行い、第2段階では、地域間比較結果を指標間で比較可能な指標として総合aggregateし、第3段階で、1個または複数個の総合指標を再び地域間で比較するものである。これが「社会指標型CPM」で、本稿1の分類のタイプ I と II が相当する<sup>14</sup>。

本稿の以下では「自治体ベンチマーキング型 CPM」における比較可能性を検討するが、「社会 指標型CPM」(国単位のものを含む)における比 較可能性と多くの論点を共有することにも着目し

図1 CPMにおける2種類のクロスセクション比較 (a)指標値の地域間比較



(b) 地域間比較結果の指標間比較



て議論を進めることとする。

CPMにおいて求められる比較可能性とは何か。比較の第1段階における地域間比較において、数字として意味のある比較可能性には、次の3段階があるといえる- (a)大小関係を比較できる(序数尺度)、(b)間隔によって比較できる(間隔尺度)、(c)倍率によって比較できる(比例尺度)。本稿の以下では、(a)を「序数尺度可測性」(厚生経済学・公共選択論の関数について序数的と呼ばれる条件に相当する)、(b)を「間隔尺度可測性」(同じく基数的と呼ばれる条件に相当)、(c)を「比例尺度可測性」と呼ぶい。情報の豊かさ、比較結果の取り扱いの容易さなど、(c)(b)(a)の順に望ましいことはいうまでもない。

CPMの第1段階の比較とは専ら地域間の比較であり、地域単位の指標値のデータ集合が、どのような尺度水準を持つかが問題となる。即ち当該指標以外に様々な条件(例えば人口規模)を有する各地域の値が、何らかの(順序尺度以上の)尺度水準で直接意味のある比較ができるような指標を用意しなければならない。各自治体において使用される業績指標は、そのままでは直接比較できないものも多いであろう。

自治体ベンチマーキング型CPMは業績測定型評価の一環である。指標値の数字は行政プログラムのアウトプット・アウトカムとしての社会状態や住民の生活水準に関し、成否や価値の評価として、その大きさが読まれなければならない。本稿では、このような意味での尺度水準を「形式的な(見かけの)尺度水準」に対する「実体的尺度水準」と呼ぶ16。

例えば、学校で行われる100点満点のテストの点数は、形式的には問題のうちのどれだけ解答できたかを表す比例尺度(零点が明確な意味をもつ)であるが、学力の実体を測るという観点からは0~100までの目盛に当てはめた間隔尺度であろう。CPMでは例えば、様々な条件が異なる地域の実測値の大小を比べる際、プログラムのアウトカム指標として同等のものと見なせるか、あるいはある業績指標について大きな実測値域での間隔と小さな実測値域での間隔を同等のものと解釈できるか、などが問われる。形式的に比例尺度であっても、実質的に大小関係を比べられなければ順序尺

度ですらなく、その数字は単なる名義尺度の記号 に過ぎなくなってしまう(後述するように、その ようなケースは珍しくない)。

第2段階以降の比較がより重要であることが多い社会指標型CPMに比して、自治体ベンチマーキング型CPMでは、個々の個別指標の比較が明確な意味を持つことが第1段階の比較において決定的に重要である"。

自治体ベンチマーキング型CPMにおける第2段階の指標間比較とは、第1段階における様々な業績指標の自治体間比較の結果を受けて、各自治体において、他自治体との比較結果を表す指標という共通の物差しによって、地域の状態や市民の生活環境の様々な側面の水準をそれらの側面同士で、あるいはそれらの水準をアウトカムとしてもたらすこととなったプログラムのパフォーマンスをプログラム間で比較することである。

第2段階における比較可能性は通常、第1段階の比較結果、すなわち自治体の指標値の位置を表す、共通の新たな指標を計算することで確保される<sup>18</sup>。相対的な位置を絶対零点をもつ比例尺度で表すことはないため、数字として意味のある比較可能性には(a)序数尺度可測性、(b) 間隔尺度可測性の2段階があることになる。ただし、理論的には、比例尺度可測性を有する絶対的位置指標も考えられる(後述)。

この第2段階の比較でも、第1段階同様、形式上でなく、実体的にどのような尺度可測性を持つかが重要である。例えば順位や偏差値といった共通の形式であっても、それらが業績測定の指標として、例えば大小を実体的に比較できるのかが問われることとなる。

第1段階の比較において用いられる指標の実体的な尺度水準が、何らかの加工や変換によって高くなることはないから、第2段階の尺度可測性は、第1段階のそれに当然制約される。即ち第2段階の比較において実体的な序数尺度を得るためには(その比較の対象となる)第1段階の各指標の実体的な序数尺度以上が、第2段階の比較において実体的な間隔尺度を得るためには(その比較の対象となる)第1段階の各指標の実体的な間隔尺度以上が必要である。その上で、第2段階の指標を、尺度水準の維持という観点から「許容された」変

換により求めることとなる。また、後述する社会 指標型CPMで用いられる相対位置指標の中には、 定義自体が尺度水準の制約をもたらすものもある。

社会指標型CPMでは、その多くの取り組みで最終目的となる第3段階の地域間比較に用いる総合指標の算出"のために欠かせない段階として、第2段階の比較可能性が位置づけられる。比較可能な指標でなければ総合化できない(欧米の文献ではしばしば「リンゴと洋ナシ(またはオレンジ)は足せない」と表現される)からである。なお、求められる比較可能性は、自治体ベンチマーキング型CPMと同一である。

このような、第1段階で用いた指標から共通の相対位置指標を算出する過程は、社会指標型CPM(国単位のものを含む)を論じる多くの文献では、標準化、規準化または基準化と呼ばれている。しかし「標準化、規準化、基準化」という語の意味や使い分けは統一されていないのが現状であり、上記を含め、CPMの方法に関する3種類の主要な操作が標準化(または規準化、基準化)と呼ばれうる - (ア)第1段階の比較のために、個々の指標を比較可能な形に加工・変換すること、(イ)第2段階の比較のために位置指標を算出すること、(ウ)(イ)の1つの方法としての標準得点化(または偏差値化)。本稿の以下では、(ア)を共通尺度化、(イ)を標準化、(ウ)を基準化と呼ぶこととする20。

第3段階の比較は、自治体ベンチマーキング型 CPMでは通常総合指標を算出しないため直接には行われないが、第2段階で用いた比較可能な相対位置指標群を用いて、総合的に比較することには意味があるだろう。すなわち総合化した指標で自治体を1次元で比較するのではなく、主成分分析やクラスター分析などの多変量解析手法により、自治体の置かれている状態の特徴や自治体間の関係を多次元で記述・観察することである。

なお、クロスセクション比較を目的とする CPMにおいても時系列方向の比較が行われる場合がある。1つは、業績指標の変化率を1つの指標 として用いる場合(後述)である。もう1つは、 自治体・地域間のクロスセクション比較の結果で ある相対的な位置関係を表す指標の時系列変化を 観察する、というケースである。後者の場合、指 標値の水準は他団体の状況に左右されるため、業績を評価するという方向での順序尺度になっていない(相対的な順位の変化はわかっても、業績自体の上昇・下降とは必ずしも一致しない)ことに注意が必要である<sup>21</sup>。

次節では、上の構図を念頭に置きつつ、指標の 比較可能性を得るために、あるいはその比較可能 性を高めるために用いられる方法について、順に 検討を加えることとする。

#### 3. 自治体間で指標を比較する方法ーその 課題と可能性

#### (1) 団体間・地域間比較(第1段階の比較)

#### ①比率の指標

個々の指標を団体間・地域間で比較可能にするための最も簡明な方法が比率の形への加工である。このような比率は、i)構成比率、ii)対立比率(異種の量と対比させて相対的な大小を表す)、iii)発生比率(分子が分母から発生する関係を表す)、iv)変化率の4種類に分類できるが<sup>22</sup>、何れもが自治体間比較で広く用いられる。構成比率は、個々の自治体における業績指標としても用いられるが、自治体間比較でもそのまま比例尺度可測性をもつ便利な比率であり、全体の中に占める割合で表せる事象は、比較が容易である。変化率も比例尺度可測性をもち、多くの指標の比較に使える。水準の比較を補う役割と、変化自体が水準よりも比較すべき業績である場合(行革による職員の削減率など)があろう。

対立比率・発生比率はあわせて関係比率と呼ばれる。地域間比較においては、分子の量(比較対象指標)の比較可能性を確保するために、比較対象指標と直接・間接の因果関係・相関関係を有することで地域ごとの(有利・不利などの)条件の差を表すような異種の量(指標)を分母とすることによって、客観的・公平な比較を可能にすることとなる。言うまでもなく、組み合わせる分母の量として適切なものを選ぶか否かが実体的な比較可能性の成否を左右する。

例えば、最近複数の文献で、道路密度(道路延 長:面積)という指標値が国際比較で突出し、日 本の道路は造り過ぎであることが分かるという趣旨の記述が見られる。例えば「国土面積 (km²)当たり全道路延長 (km)」は米0.95、加0.15、仏1.73、英1.60に対し日本は3.16、これを可住地面積当たりで再計算すると米2.00、加0.33、仏2.80、英2.48に対し日本は14.67であるという (五十嵐・小川2008)。

しかしここで別の関係比率として「人口(千人) 当たり」を計算すると米32.5、加44.6、仏16.2、 英6.6、日9.3に、「自動車保有台数(千台)当たり」 は米37.2、加69.9、仏25.6、英12.0、日16.2となる (人口、自動車保有台数は総務省『世界の統計 2011』に掲載されている数字を使用)。突出して いるのは道路ではなく、狭い国土に人口や車が桁 外れに密集し、道路は一側面に過ぎないことが分 かる<sup>23</sup>。

#### ②関係比率による指標の共通尺度化

自治体ベンチマーキング型CPMにおいて、意味のある量と組み合わせて関係比率を作ることはほとんどの場合、自治体の規模の差(人口やサービス対象人口の差など)を調整して、指標の共通尺度化を図るものである。そこでは、共通尺度化した指標aについて次のような関係を想定していることになる。

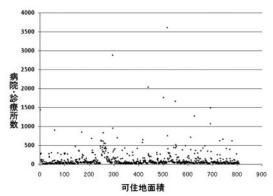
#### 「元の指標値」 = $a \times$ 「分母に用いる量」 (1)

つまり、分母に用いる量と元の指標値が概ね (本質的に) 比例関係にあり、aの大きさこそが、 当該自治体固有の状態を示す。

さて、しばしば実際のCPMに用いられている 関係比率の形の指標に「可住地面積当たり病院・ 診療所数」という指標がある(例えば週刊ダイヤ モンド編集部 2007)。ところが全国市区の可住地 面積と病院・診療所数の関係を散布図に描くと図 2のようになる。

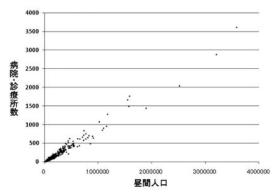
図2をみれば、可住地面積との相関が全くないことは明らかであり、この比率の意味が曖昧であることがわかる。一方、昼間人口<sup>24</sup>と病院・診療所数の組合せで散布図を描いたのが図3である。両者が概ね比例関係にあり、この組合せで関係比率を作れば意味のある共通尺度が得られることが

図2 病院・診療所数と可住地面積の関係



(出所) 総務省『統計でみる市区町村のすがた2010』掲載の市 区データより筆者作成。

図3 病院・診療所数と昼間人口の関係

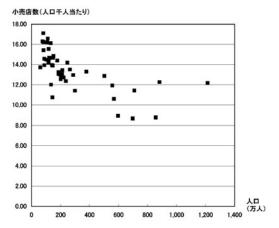


(出所) 図2に同じ。

わかる。可住地面積当たり病院・診療所数という 指標では医療環境の比較は全くできず、可住地人 口密度に近い指標をそれと知らずに用いているに 等しい。

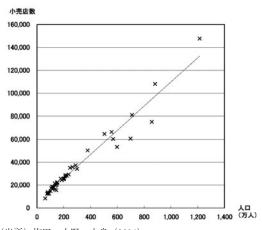
一般に経済・社会の活動に関わる多くの指標は 人口規模との関係が想定されるが、その関係は上 記のように概ね比例関係にあると見なせるとは限 らない。図4は都道府県別の人口千人当たり小売 店数を、図5は小売店数の実数を人口に対してプロットしたものである。図4からは、人口が少ない県で指標値が高い傾向がはっきり見て取れる。 その理由は図5より明らかで、概ね直線状に分布しているものの、その直線は原点を通らない。即 ち比例関係(原点を通る直線)を想定した共通尺 度では、人口の効果を調整することはできない。

図4 人口当たり小売店数と人口の関係



(出所) 梅田・小野・中泉 (2004)

図5 小売店数と人口の関係



(出所) 梅田・小野・中泉 (2004)

より一般に、地域の経済・社会に係る多くの指 標値と人口との関係は(1)式のようにはならず、 いわゆる規模の効果(効用逓減、効用逓増)・人 口弾力性によって、指数関数的になることが知ら れている(日下部2002、2003・2004)。人口(昼 間人口、年齢階級別人口などを含む)その他の量 の規模を踏まえた共通尺度化を図る場合、(1) 式 からの乖離が実務上無視できないようであれば、 関係比率は比較可能な共通尺度にならず、以下の ④で述べる回帰分析などによる調整が必要とな る。

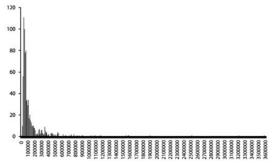
#### ③変数変換、とくに対数変換

図6・7に全国の市区の昼間人口及び対数変換し た昼間人口のヒストグラムを示す。

このように対数変換(実数との対応が容易にな るよう10を底としている)すると自然な分布25に なるデータの背後には、量が等比(数列)的、指 数関数的に成長する過程がある。人口は、等比的 に成長するのが自然であることが直感的に納得で きる例だが、世の中の複雑な過程によって産み出 される数・量の多くがそうであることが知られて いる260

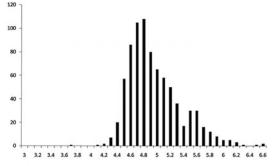
等差的でなく等比的に成長する過程から生み出 される数・量(例えば人口)は、形式的には比例 尺度であっても、小さな値域と大きな値域におい てその間隔が実質的に同じ意味を常に持つか否か は多くの場合疑問である。このことは、自治体の 業績測定型評価において、等比的な性格を持つと 考えられる多くのアウトカム指標の将来の目標値 を設定したり、一定の時間経過後に目標達成度を

図6 昼間人口の分布(市区)



(出所) 2005年国勢調査結果より筆者作成

昼間人口の対数の分布(市区)



(出所) 2005年国勢調査結果より筆者作成

評価したりする際に注意を要するが(小野2008、2010)、幸いCPMでは、②で共通尺度化した指標として取り上げた「昼間人口当たり病院・診療所数」のように、等比的な性格を持つ変数(規模の大きさを表す変数は多くが該当するだろう)で除す場合、この問題をある程度回避できる"。

ただその場合でも、元の指標データが本来等比 的な性格を持つことは認識すべきであり、関係比 率による共通尺度を検討する際には、原数値では なく対数変換した値で散布図をみる(あるいは両 対数紙にプロットする)ことが望ましい。

また、②で言及したような、指標値と人口の関係が単純な比例(直線)関係からどの程度ずれ、いわゆる規模の効果があるか否かは、両方を対数変換して、回帰直線の傾きが1に近いか否かによって判断することができる(1より大きければ効用逓増、など)(日下部 2003・2004)。

もちろん、指標データの背後にある構造次第では、対数変換ではなく、逆数変換や平方根変換などによって自然な分布となる可能性もある。評価において実体的な意味をもつ間隔尺度であるためには、指標データが等差的な線形構造に対応する自然な分布を持つ必要があることを忘れるべきでないだろう。

#### ④説明変数による調整一調整済み業績指標 (APMs)

調整済み業績指標adjusted performance measures (以下APMsとする)とは、個々の組織がコントロールできない(あるいは難しい)が業績指標の値に影響を及ぼすような変数・要因について調整するもので、具体的にはそのような変数で業績指標の値を回帰した残差、すなわち回帰式28が予想する値と実績値の差を指標とする(Stiefel et al. 1999)。

団体iの業績指標 $P_i$ を、説明変数 $Z_i$ で次のように回帰したとすると( $b_0$ は切片、 $b_1$ は回帰係数、eは誤差項)、

$$P_i = b_0 + b_1 Z_i + e_i \tag{2}$$

この指標P.のAPMは次の残差で表される。

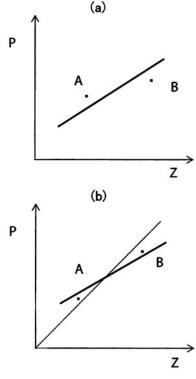
$$M_i = P_i - b_0 - b_1 Z_i$$
 (3)

図8 (a) に2団体A、Bの業績Pを比較するイメージを示す。変数Zによる回帰直線からの距離によって比較すれば、業績の大小関係が逆転することも (2団体を含む全団体の順位が大いに入れ替わることも) 大いにあり得る。

図8 (b)は変数Zを分母とする関係比率(図中の対角線との位置関係に相当)では実測値と大小関係が逆転するが、APMsでは再度逆転するケースである。②で言及した人口による調整の例など、このような可能性もある<sup>29</sup>。

自治体ベンチマーキング型CPMに登場する多くの業績指標は、上で言及した人口の他にも、サービス対象顧客の属性、天候・気象など行政がコントロールできない様々な要因の影響を受ける。個々の団体内の業績測定であればよいが、自治体間で比較するとなれば公平を欠くこととなる。自治体ベンチマーキング型CPMの取り組みが多い米国では、業績指標の差を説明する分析として回帰分析が行われ、APMsが公平な指標として用い

図8 APMsによる比較のイメージ



(出所) 筆者作成

られている(Kopczynski & Lombardo 1999, Rubenstein et al. 2003)。

表1は、ICMA主導のCPMの実例からの抜粋である。ここでは図書館利用指標を世帯収入で調整しているが、この表及び図(回帰直線からの距離が視覚的にわかるグラフ、本稿では省略)の下の注では、この他に影響を与える可能性がある要因として、図書館の支所数や開館時間、図書館の設備、サービス対象人口の経済的・人口学的特性を挙げている。

このように影響を与える変数が複数ある場合、複数の変数で業績指標を説明する重回帰分析を行うことも考えられるが、変数選択をしなければならない。統計的に有意な回帰係数をもつ変数を回帰モデルに残すのが基本となるが、変数間に一定の相関がある場合など、必ずしも容易な作業とはならない。

業績指標に影響を与える変数としては、当該自 治体がコントロールできない外部要因のほかに、 例えば評価対象プログラムに投入する職員数な ど、当局がコントロールできる様々な内部要因も 考えられる。表1に絡んで言及した図書館の支所 数や開館時間なども、関連政策・事業を通じてコ ントロールできる変数である。このような内部要 因を説明変数としたAPMsは、各団体の効率を表 す指標であり比較・分析の対象として有用である が、もはや住民の享受するアウトカムそのものと はいえない。自治体間でパフォーマンスを比較す るためには、外部要因で調整するのが基本であろ う。

なお、APMsは回帰分析であるから、回帰直線が意味を持つこと、すなわちデータの分布によく当てはまることが前提となる。計算結果を使う前には、回帰モデルの説明力が十分か否かを決定係

表1 1人当たり図書館利用の世帯収入による調整例

団体	図書館利用	世帯収入	回帰による予想
(報告数15)	(1人当たり)	(中央値)	値からの距離
シンシナチ	19.4	21,006	15.1
プリンスウィリアム郡	36.4	56,828	9.3
タスコン	11.7	21,748	6.9
:	:		
リッチモンド	6.3	36,989	-8.2

(出所) Kopczynski & Lombardo (1999) に掲載の図より抜粋。

数(説明変数が指標値の変動の何%を説明できたかを示す)などにより判断する必要がある。残念ながらこの数字に万能の基準はなく、誰かが判断をしなければならない。Morley et al. (2001) によれば、ICMAのプロジェクトでは決定係数15%を判断基準としている。

ところで、表1に示したICMAプロジェクトの例ではAPMs即ち残差の値(回帰直線からの距離)の降順に団体が並んでいる。このランキングはしかし、世帯収入に応じて図書館利用が増大するという関係を踏まえると必ずしも公平とはいえない。同一の条件下でどれだけ図書館が利用されているかを比べようとすれば、関係比率の共通尺度として、予想値(標準ライン)からの乖離の比率を比較すべきであろう。表1中のランキングが2位のプリンスウイリアム郡が指標値の26%相当分標準を上回っているのに対し、3位のタスコンは59%相当上回っている。タスコンのパフォーマンスの方が優れているというべきであろう。

因みに統計学では回帰モデルの診断をするためにしばしば残差の分布を観察するが、その際、残差をその標準偏差の推定量で割る基準化30を行うことが多い。しかし、この基準化は第一義的にモデル間の比較を可能にするためのものである。またその基準化の中で二義的に、各データの回帰直線への影響力を踏まえた補正が残差のそれぞれに加えられるが、実績値としての残差に着目するCPMでは比較可能性を高めるものとは考えられない31。

また、通常の(線形)回帰分析は、変数間の線形の関係を前提としている(正規分布を仮定できればなお都合がよい)。したがって、回帰分析の前に、実測値の分布によっては本節の③で述べたような変数変換を行う必要がある。例えば、もとの業績指標のデータと外部要因である説明変数のデータの一方または両方が等比的・指数関数的な性格を持つのであれば、対数変換を行うことで決定係数の高い回帰直線が得られる。つまり比較可能性の高いAPMsが得られることとなる。

# (2) 標準化のための位置指標と指標間比較 (第2 段階の比較)

#### ①位置を表す指標への標準化

自治体ベンチマーキング型CPMにおける第2段階の比較、すなわち個々の自治体において、指標間・プログラム間の(地域間比較結果の)比較を可能にする方法は、社会指標型CPM(国単位のものを含む)において総合指標を算出するために不可欠な段階である、指標の標準化とほぼ共通である。総合化するためには各指標を全く同列に扱う必要があるからである(ウェイトは別途付与するとして)。

この標準化とは、すでに述べた通り、共通の位置指標を算出することに他ならない。指標・プログラムを積極的に比較する意図がなくても、個別の指標ごとに当該自治体の位置を明確にする位置指標を算出すれば、それは自動的に、共通の位置指標を持つプログラムの間で比較をすることになる。

一般に、指標を標準化するための位置指標には 絶対的位置指標と相対的位置指標がある。

#### ①-1 絶対的位置指標32

標準化のために、絶対的位置を明らかにする手法として提案されたのが、各指標について規範的な基準点を設け、それとの関係で各自治体の絶対的位置を示す基準点方式である。これらの絶対的位置指標を実際に算出できれば、相対的な位置も直ちに明らかになる。

基準点方式は、基準点の数によって、1基準点方式、2基準点方式、3基準点方式などと呼ばれる。1基準点方式とは、何らかの方法によって誰かが定めた、例えば「望ましい」水準と実測値の比や差を指標化するものである。2基準点方式は、例えば指標iについて「最低限の充足水準」 $X_{iM}$ と「十分に充足された水準」 $X_{iF}$ を定め、地域jの実測値 $x_{if}$ の絶対位置指標 $y_{if}$ を次のように算出する。

$$y_{ij} = \frac{x_{ij} - X_{iM}}{X_{iF} - X_{iM}} \tag{4}$$

3基準点方式は、おそらくは基準点方式として 最も早くドレヴノフスキが提唱したもので、上記 の2基準の他にニーズが全く充足されていない 「ゼロ点」も定め、実測値の値によって場合分け して、絶対位置指標に変換する (Drewnowski 1974)。

ここで、絶対的位置指標によって指標間比較をするためには、「最低限」や「十分」がどの指標についても同一の「程度」を意味するという、文字通りの絶対性が求められる。かつて基準点方式の指標を作成した研究では、専門家へのアンケートを行うなどの策が講じられたが(河野 1985)、このような基準点の設定において任意性・恣意性を排除することは現実的でなく、実務上の継続的活用を想定する自治体ベンチマーキング型CPMにおいて、採用は難しい。

#### ①-2 相対的位置指標

各指標について、各自治体の全自治体における相対的な位置を表す指標である。絶対的位置指標とは異なり、実測値データから計算される。社会指標型CPM(国単位のものを含む)などでは、相対的位置指標として次のようなものが用いられている<sup>33</sup>。

A. 順位

$$y_{ij} = Rank(x_{ij}) \tag{5}$$

B. カテゴライズ (多段階評定)

$$y_{ij} = 1, \dots, k \tag{6}$$

(*k*はカテゴリー数)

C. 基準点(平均または中央値)との比

$$y_{ij} = \frac{x_{ij}}{x_{i}}. (7)$$

 $(x_i.$  は指標 i の平均または中央値)

D. 基準点 (最大値) との比

$$y_{ij} = \frac{x_{ij}}{\max_{j}(x_{ij})} \tag{8}$$

E. 基準点 (参照国・地域) との比・差

$$y_{ij} = \frac{x_{ij}}{x_{ir}} \text{ or } y_{ij} = \frac{x_{ij} - x_{ir}}{x_{ir}}$$
 (9)

(rは参照国・地域)

F. Min-Max (最大值最小值間相対位置)

$$y_{ij} = \frac{x_{ij} - \min_{j}(x_{ij})}{\max_{j}(x_{ij}) - \min_{j}(x_{ij})}$$
(10)

または

$$y_{ij} = \frac{\min_{j}(x_{ij}) - x_{ij}}{\max_{j}(x_{ij}) - \min_{j}(x_{ij})}$$
(11)

G. 標準得点 (zスコア)、基準化

$$y_{ij} = \frac{x_{ij} - x_i}{s_i} \tag{12}$$

(s,は指標iの標準偏差)

A、Bは順序尺度の指標であるため、間隔尺度 とみなさない限り、形式的にも大小の比較しかで きない。また実測値の持つ情報の大半を捨てるこ ととなる。本来、順位や評点は、他の相対的位置 指標を補う役割であろう。

C~Eは、実測値データのごく一部しか使っておらず、C・Eは外れ値などに対して頑健という面はあるが、感度の低い指標である。例えば、基準として参照すべき地域が明確であるなど、特別な要請がない限り、使用する利点はない。

Fの"Min-Max"は、世界的に広く知られる人間開発指数Human Development Indexなど、多くの社会指標の標準化に使われている。分母が分布の「範囲(レンジ)」であるため外れ値の影響を受けやすく、CPMで継続的に観察する際に、最大値・最小値の変化が大きい場合、全地域の計算結果が大きく様変わりしてしまうことになる。一方、取りうる値の範囲が0~1と分かりやすく、計算式が極めてシンプルであることは、Gの標準得点と比べて大きなメリットである。

実測値の集合を平均ゼロ、分散(標準偏差)1のデータに変換するGは、社会指標の標準化で最も多く使われる(Freudenberg 2003)だけでなく、統計解析の有りと有らゆる場面で行われる基準化そのものである<sup>34</sup>。

標準得点とMin-Maxの大きな違いは、前者が実 測値の分布に応じて変換後の値域が変わるのに対 し、後者は実測値がどのような分布であっても変 換後の値が0~1の範囲に収まることである。また 後者では、個々の実測値を変換した値が分布の形 に関わりなく、つまり最大値・最小値以外の地域の実測値に関わりなく決まる。これらのことから、少なくとも自治体ベンチマーキング型CPMの相対的位置指標として、標準得点のほうが比較可能性において優れていることは明らかである。

#### ②偏差值

上の標準得点 (Zスコア) を、次式のように線 形変換して、平均50、標準偏差10としたものが偏 差値である。間隔尺度の標準得点を線形変換した ものであり、相対的位置指標として、両者は全く 同等の情報をもつ。

$$Z_{ij} = 50 + 10 \times \left( \frac{x_{ij} - x_i}{s_i} \right) \tag{13}$$

偏差値は平均が50点で、百点満点の半分に相当するようにも感じられ、多くのデータが平均点のまわりに「適度に」散らばることが多いのがミソである。そして学校教育において広く使われた結果、一定の年齢以下の日本人なら誰でも直感的に数字の大きさを理解できることも、偏差値を使うことの現実的なメリットである。

偏差値は形式的には、実測値が間隔尺度以上のデータであれば計算できる。しかし、実体的な間隔尺度可測性がないと偏差値の意味は曖昧になる。先にみたように等比的・指数関数的構造をもつデータでは、大きな実測値になるほど偏差値は大きく歪み、突拍子もない値を取ることもある(注24参照)。実体的な間隔尺度可測性がなければ、必要な変数変換(上の場合は対数変換)を施してから偏差値を計算すればよい。。

一方、偏差値は本来、正規分布しているデータの場合に正しい意味をもつ、という趣旨の説明がしばしばなされるが、これは必ずしも正しくない。試験の得点データは正規分布に近い分布になることが多く、偏差値の値を正規分布における位置に対応させる(偏差値60は上位から16%に相当など)習慣があるというだけである。

ただし、正規分布しているデータは、実体的にも間隔尺度以上であると考えられるから、正規分布は偏差値使用の十分条件ではある。必要条件ではないから、正規分布から離れた分布で計算した偏差値は意味がないという言説は誤りである36。

実体的に間隔尺度以上であって、平均と標準偏差が明確に定まる限り、分布のピークが2つあろうが、右・左に歪んでいようが偏差値は常に明確な意味をもつ。

なお、計算された偏差値は間隔尺度のデータであるから、倍率で比べたり増減率を計算したりすることはできない。

# 4. 自治体ベンチマーキングの今後に向けて

自治体ベンチマーキングの基本は数字の比較である。適切な比較は有意義であり、不適切な比較は有害である。

適切な比較のために統計学的取扱いが求められる場面もある。例えば回帰分析によってAPMsを計算する場合、指標の比較可能性を高めようとすれば、その分よりオーソドックスな手続きや厳密な取扱いが求められることとなる(回帰分析を機械的・形式的に行うことは極めて危険である)。一方で、自治体ベンチマーキング型CPMに分かり易さが求められるも当然のことである(肝心の部分がブラックボックスでは結果がシンプルでも分かり易いとはいえない)。このことは、自治体ベンチマーキング型CPMの実務において折り合いをつけねばならぬ悩ましい課題である³³。

地方分権の進展に応じた経営戦略策定、行政経 営のさらなる効率化、人口減少や財政逼迫に伴う サービス提供量見直しなど、自治体ベンチマーキ ング情報が活用されるべき場面は増えるであろ う。各方面での体制整備やノウハウの蓄積、議論 の深化が望まれる。とりわけ本稿で論じたような 指標の比較に係る基本については、評価論・統計 学などの分野の各種リソースの充実と活用が求め られる。

#### 謝辞

本論文の初稿に対して匿名の査読者から、論旨 を明確にする上で大変有益な指摘をいただいた。 深く感謝します。

#### 注記

- 1 比較しなくても意味が明確な指標の例として、各種 サービスの普及率や利用率、苦情やトラブルの件数 などを挙げることができるが、これらは100%、ゼ ロ件という基準と比べているともいえる。
- 2 指標の実績値を目標値と比較すること (及びその比較に基づく何らかのアクション) も「基準との比較」という意味でベンチマーキングと呼ぶことができる。しかし本稿では本特集の趣旨に沿って、ベンチマーキングという語を自治体間・地域間の比較に限定して用いる。
- 3 業績指標のクロスセクション比較には、プログラム間の比較という別の重要な局面もある。本稿では2 (2)以降で地域間比較を目的とする取り組みの枠組みの中で行われるプログラム間比較に言及する。
- 4 ここで社会指標とは「広く地域社会に関わるアウトカムを測定する指標」という意味の、業績指標の1類型としての社会指標ではない。18世紀まで遡る起源をもち、1950,60年代の西欧諸国における「経済発展の計測だけで社会の進歩を語れるか」という問題意識によって広く取り組まれる運動ともなった、「国民・住民の厚生水準を測るモノサシ」を意味する社会統計の用語としての社会指標である。指標を作成・活用する取り組み、指標の体系、個々の指標の何れの意味でも用いられる。

なお、社会指標は厚生経済学・公共選択論における効用関数・社会厚生関数と接続して論じられることがある。そもそも個人の関数が可測か否かという議論から始まる効用関数と、指標の実測値を前提とする社会統計としての社会指標は「著しく異なる」(Drewnowski 1974) 性格をもつが、ともに社会・生活の状態・水準を測定する装置を巡る議論であり、本稿でも一部参考にしている。

- 5 筆者は旧経済企画庁のPLIを拡張する試みとして都 道府県別の「豊かさ指標」をかつて提案した。小 野・田渕(2001)第6章、三菱総合研究所(1997) (1998)を参照のこと。
- 6 これら2つの事例について比較分析した論文にCoe (1999)、Ammons et al. (2001) などがある (後者は 1991年開始のInnovation Groups Projectも加えた比較)。
- 7 本特集に掲載の熊倉 (2011) を参照のこと。

- 8 本特集に掲載の中村 (2011) を参照のこと。
- 9 この他、筆者がかつて関わった、小規模ではあるが 比較的早い時期(2000~2001年度)の事例に、鎌倉 市・逗子市・藤沢市が参画した「湘南ベンチマーク」 がある。逗子市HPに「平成12年度湘南ベンチマー クについて」などの情報が掲載されている。
- 10 数学では、ある集合(半順序をもつ集合)の元a、b についてa≦bまたはb≦aが成り立つとき、aとbは比 較可能comparableであるという(任意の元の組合せ について比較可能であれば一即ち完備律を満たせ ば、全順序集合となる)。これはCPM・社会指標の 比較可能性の観点からは、そもそも数字が比較でき るか、すなわち大小関係があるとみなせるか否かと いう最も入口の条件に相当する(本文2(2)参照)。
- 11 欧米の社会指標の文献においても本稿と同様、比較 可能性という意味でcomparabilityという用語が広く 使われるが、その他にcomparabilityに近い意味で用 いられている語にuniversality及びcompatibilityがある (本稿が参照した文献ではUNCTAD (2004)、 Kalimo (2005)、Stiglitz et al. (2009) など)。

なお、comparabilityが限定的に使われる場合もある。社会選択論のアプローチで社会指標を論じたWelsch(2005)では、1変数に係る条件をmeasurability、複数変数に係る条件をcomparabilityと規定している。また比較可能性の構成概念の一部である「同じ物差しで測られているか」という概念について、commensurable (Ebert and Welsch 2004)、commensurate (Nardo et al. 2005)(尺度共通性、単位共通性)という語も用いられる(本稿2(2)でいう「共通尺度化」とは異なる)。

- 12 妥当性(測るべきものを測ること)と信頼性(同じ条件から同じ結果が、あるいは同じ事象の測定から同じ結果が得られること)は、評価・業績測定に留まらず、社会科学における様々な測定における原理であり、求められる条件である。業績指標の妥当性・信頼性については、例えば小野(2009)を参照。
- 13 このような指標の妥当性の分類については、梅田・小野・中泉 (2004)、小野 (2009) を参照のこと。なお、社会指標を政策指標として論じた初期の代表的文献である "Policy Indicators" (MacRae 1985) は、社会指標の検討においても、社会科学一般と同様conceptとmeasureの区別すなわち概念と指標(測定する具体的手段=物差しとしての指標)の区別が最

- 初に求められると強調している。"Policy Indicators" のいう指標の質は、本稿でいう第2段階の妥当性と 測定の信頼性とに左右されることとなる。
- 14 タイプ  $I \sim \mathbb{N}$  のCPMのこの2類型への再分類には、タイプ  $I \cdot \mathbb{I}$  で指標の総合化を行わないなど、若干の例外もあるかもしれない。
- 15 Ebert and Welsch(2004)、Welsch(2005) における interval-scale measurability、ratio-scale measurabilityの 用語に依拠する。
- 16 筆者が調べた限りでは、このような区別に対応する、 定着した用語はない模様である。本稿では、竹内 (1971) における「量の実体化」という概念を参考 に、実体的な尺度水準と呼ぶことにした。
- 17 社会指標型CPMでは、例えば個々の実測値によって各地域の厚生水準を直接読めない場合でも、互いに関係する複数の指標を総合することで、その総合指標の値が明確に意味をもつようにできる可能性がある。
- 18 実際には、共通の指標を計算できる指標間でのみ比較ができることになる。
- 19 Stiglitz et al. (2009) は、社会指標における総合指標 の算出を、経済発展の指標を超えるためのあらゆる 努力の"holy grail" (キリストが最後の晩餐で用い た聖杯。アーサー王伝説において探求される)にも 譬えられるとしている。ただし、総合指標の算出の 是非に関しては尽きない論争があり、また総合指標 を算出するために各指標に付与するウェイトの算出 方法を巡る議論もまた尽きない。ともに最終的には、 指標作成者の意思と判断に任される部分が残るとい える。これらの課題については、社会指標創成期の 文献 (例えばDrewnowski (1974) など) から、近 年の社会指標リバイバルの世界的気運のなか各機関 において総括的に社会指標の方法を論じた文献(例 えば仏サルコジ大統領のイニシアティブによる Stiglitz et al. (2009)、OECDとECによるOECD& European Commission-JRC (2008) など) に至るま で、多くの文献において詳細に取り上げられている。
- 20 英語文献では、(a)~ (c)の操作は、standardization またはnormalization (欧州の文献では英国流綴りの standardisation, normalisation) と呼ばれる。(c)を除いて、変数を変換するという意味でtransformationと いう語も用いられる。これらの語の使い分けは日本 同様不統一で、本稿が参照している文献でも使い分

けは区々である。

- 21 例外は、3(2)①で紹介するドレヴノフスキの3基準 点方式である。絶対的な基準となる水準を定めるこ とによって、相対的な位置とともに個々の地域の時 系列変化も直接読める。しかし本文で後述する通り、 絶対的な基準を定めることの任意性・恣意性が実務 上決定的な制約となる。
- 22 宮川 (2003) による整理を参考にしている。
- 23 日本の道路が造り過ぎであるか否かについては、全 道路延長の国際比較ではなく、道路建設コストの関 係比率による国際比較や、地域別の偏在(道路が足 りない地域で造る余地がなく、多すぎる地域で建設 が続く)を表す指標の国内比較が、端的にして有用 な指標となろう。
- 24 通勤通学による人口移動を考えれば、都市部を中心に量的な医療環境が常住人口ではなく昼間人口に左右されることは明らかであろう。なお、本稿で言及しているこの都市ランキングでは、犯罪発生率など複数の指標を昼間人口ではなく登録人口により「共通尺度化」している結果、東京都心部の区などでは260や-90など極端な偏差値が頻出し、結果全体を大きく歪めている。
- 25 統計学において変数変換は、変換後に図7のような 簡潔な線形性や正規性を達成して統計分析手法の適 用範囲を拡げるために行われる。図7の分布は、そ の意味で「自然」である。
- 26 1996年に証明されたベンフォードの法則(世の中の自然発生的なデータの先頭桁の数字がdとなる確率はlog<sub>10</sub>(1+1/d)。即ち確率は均一ではなく、「1」の確率が一番高い)がこのことを傍証している。世の中の多くの量がこの法則に従うことが知られているが、それは等比的に成長することを意味する。
- 27 本文②において共通尺度化し損ねた指標の例として 取り上げた「可住地面積当たり病院・診療所数」は、 図2からも等比的なパターンの分布となることが明 らかである。これは、実体的に間隔尺度とはみなせ ない指標について、間隔尺度以上が計算の前提であ る偏差値を計算している例でもある。
- 28 本稿では回帰モデルとして、専ら1次式で表される 線形回帰(直線回帰)モデルを想定しているが、理 論的には非線形回帰(曲線回帰)モデルの適用もあ りうる。もっとも、目的変数である自治体の業績指 標と何らかの説明変数の間に(変数変換を施しても

- なお)曲線関係を仮定すべきケースは多くないと思われる。またCPMにおけるAPMs の算出においては、いわゆるパーシモニー(節約)の原理に従うべきで、回帰モデルはシンプルな方が望ましいであろう。
- 29 図8 (a) (b) では作図上、回帰曲線を挟んで反対側に 2団体の点を描いているが、APMsはあくまでも回 帰直線からの距離である。回帰直線の同じ側にある 点同士でも、同様のことが起こる。
- 30 規準化、標準化、(内的) スチューデント化とも呼ばれる。本稿では「基準化」とする。
- 31 残差の基準化によって回帰モデル間の比較可能性が高まることは、APMsにおいて業績指標間(プログラム間)の比較可能性に関わるが、どの業績指標にも同じ説明変数を使わない限り特に意味はない。また日下部(2002、2003・2004)は、基準化によって起こる個々の残差の補正に積極的意味を与え、地域の統計指標による比較は基準化残差を用いるべきであると主張している。しかし、APMsの比較可能性を高める上では、本文で述べたように、比例尺度の業績データの残差を予想値との比率で比較することの方が大きな意味を持つと考えられる。
- 32 一部、三重野 (2010) による整理に依拠している。
- 33 このリストの作成にあたっては、Freudenberg(2003)、 Jacobs et al. (2004)、Nardo et al. (2005)、OECD& EC-JRC (2008) を参考にしている。
- 34 本稿1で言及した米国の社会指標型CPM "Places Rated Almanac"のデータを統計学的に分析した Becker et al. (1987) は、実測値に対して対数変換・逆数変換・平方根変換を適宜施した後、全指標を基準化して比較可能性を高めた"common-sense version of the data"としてから、分析を進めている。
- 35 自治体ベンチマーキング型CPMで、実測値に変数 変換を施してから偏差値を計算する場合、そのよう な手続きの必要性について、各方面に少し丁寧な説 明が必要かもしれない。
- 36 偏差値とは平均50、標準偏差10の正規分布に換算したときの点数であるという、これもしばしば目にする説明もまた誤りである。もとのデータが正規分布でなければそうならない。偏差値の計算は線形変換であり、幅を除いて分布の形は変わらない。
- 37 イギリスの健康・教育分野の地域別業績指標に対し、APMsによる比較などの分析を純粋に統計学的に行ったGoldstein and Spiegelhalter (1996) の後半に

付された討議部分には、この種の業績指標を発表する実務と、統計学による厳密な分析との間のギャップの大きさをうかがわせる議論が記録されている。このようなギャップはCPMに限った問題ではなく、日本においても定量的手法に馴染みのない行政職員が業績指標を取り扱う際など、そのことを意識していない場合を含め、様々な困難に直面していると考えられる。

#### 参考文献

- 五十嵐敬喜・小川明雄 (2008) 『道路をどうするか』 岩波書店
- 岩手県総合政策室調査統計課(2001a)『平成13年度岩 手県県民生活指標』
- 岩手県総合政策室調査統計課(2001b)『平成13年度岩 手県地域統計指標-広域生活圏・地方振興局・市町 村の指標-』
- 梅田次郎・小野達也・中泉拓也 (2004) 『行政評価と 統計』、日本統計協会
- 大友篤(1997)『地域分析入門(改訂版)』、東洋経済 新報社
- 小野達也(2002)「行政の改革における社会指標の役割-行政運営におけるアウトカム測定」、『富士通総研Discussion Paper』、01-34
- 小野達也 (2008)「都道府県の評価の10年-定量評価 のこれまでとこれから」、『日本評価研究』、8(1):19-38
- 小野達也(2009)「業績測定型評価のかんどころープログラムのロジックと指標の妥当性ー」、『評価クォータリー』、11:2-17
- 小野達也 (2010) 「業績測定型評価のかんどころ (2) -目標設定・達成度評価における数・量の取扱 いー」、『評価クォータリー』、14:2-17
- 小野達也・田渕雪子 (2001)『行政評価ハンドブック』、 東洋経済新報社
- 河野博文 (1985)「社会指標の現状と理論的諸問題」、 『季刊・社会保障研究』、21(1):72-80
- 日下部眞一(2002)「回帰偏差値の考案とその効用: 地域間格差を相対評価する偏差値」『広島大学総合 科学部紀要、理系編』、28:109-126
- 日下部眞一  $(2003 \cdot 2004)$  「地域経済格差の正しい理解のために $(1)\sim(3)$ 」、『統計』、2003年11月号、12月号、2004年2月号

- 熊倉浩靖(2011)「都市行政評価ネットワーク会議: 自治体業務改善のためのベンチマーキング」、『日本 評価研究』、11(2):31-44
- 週刊ダイヤモンド編集部 (2007)「全国805都市ランキング:安心して住める街」、『週刊ダイヤモンド』、8/11・18合併号:30-72
- 竹内啓 (1971)『社会科学における数と量』、東京大学 出版会
- 中村葉子 (2011) 「福井市の行政評価と自治体ベンチマークシステム」、『日本評価研究』、11(2):55-64
- 三重野卓(2010)『福祉政策の社会学-共生システム 論への計量分析-』、ミネルヴァ書房
- 三菱総合研究所(1997)『生活の豊かさと満足度の分析-主観的評価を踏まえた新たな地域の豊かさ指標の試み-』
- 三菱総合研究所 (1998) 『生活の豊かさと満足度の分析-三菱総研「豊かさ指標」98年版-』
- 宮川公男 (2003) 『統計学でリスクと向き合う』、東洋 経済新報社
- Ammons, D.N., Coe, C. and Lombardo, M.(2001).

  Performance-Comparison Projects in Local Government:

  Participants' Perspectives. *Public Administration Review*,
  61(1), 100-110.
- Becker, R. A., Denby, L., McGill, R. and Wilks, A. R. (1987). Analysis of Data From the Places Rated Almanac. *The American Statistician*, 41(3), 169-186
- Coe, C. (1999). Local Government Benchmarking: Lessons from Two Major Multigovernment Efforts. *Public Administration Review*, 59(2), 110-123.
- Drewnowski, J. (1974). On Measuring and Planning the Quality of Life. The Hague:Mouton. (=阪本靖郎訳 (1977)『福祉の測定と計画』日本評論社)
- Ebert, U. and Welsch, H. (2004). Meaningful Environmental Indices: A Social Choice Approach. *Journal of Environmental Economics and Management*, 47, 270-283.
- Freudenberg, M. (2003). Composite Indicators of Country Performance: A Critical Assessment. DSTI/IND(2003)16, OECD Paris.
- Goldstein, H. and Spiegelhalter, D. J. (1996). League Tables and Their Limitations: Statistical Issues in Comparisons of Institutional Performance. *Journal of the Royal Statistical Association*, 159, 385-443.

- Jacobs, R., Smith, P. and Goddard, M. (2004). Measuring Performance: An Examination of Composite Performance Indicators. Centre for Health Economics, University of York, Technical Paper Series 29.
- Kalimo, E. (2005). OECD Social Indicators for 2001: A Clitical Appraisal. Social Indicators Research, 70, 185-229.
- Kopczynski, M. and Lombardo, M. (1999). Comparative Performance measurement: Insights and Lessons Learned from a Consortium Effort. *Public Administration Review*, 59(2), 124-134.
- MacRae, D. (1985). Policy Indicators: Links Between Social Science and Public Debate. Chapel Hill, NC: University of North Carolina Press.
- Morley, E., Bryant, S.P. and Hatry, H.P. (2001). *Comparative Performance Measurement.* Washington, D.C.: Urban Institute.
- Nardo, M., Saisana, M., Saltelli, A. and Tarantola, S. (2005). Tools for Composite Indicators Building. EUR21682EN. European Commission-JRC:Italy.
- OECD and European Commission-JRC (2008). Handbook

- on Constructing Composite Indicators: Methodology and User Guide. Paris
- Rubenstein, R., Schwartz, A.E. and Stiefel, L. (2003). Better than Raw: A Guide to Measuring Organizational Performance with Adjusted Performance Measures. *Public Administration Review*, 63(5), 607-615.
- Stiefel, L., Rubenstein, R. and Schwartz, A.E. (1999). Using Adjusted Performance Measures for Evaluating Resource Use. *Public Budgeting & Finance*, 19(1), 67-87.
- Stiglitz, J.E., Sen, A. and Fitoussi, J-P. (2009). Report by the Commission on the Measurement of Economic Performance and Social Progress. Insee, OFCE, and OECD.
- United Nations Conference on Trade and Development (2004). Review of the Comparability and Relevance of Existing Indicators on Corporative Social Responsibility. TD/B/COM.2/ISAR/24. Geneva.
- Welsch, H. (2005). Constructing Meaningful Sustainability Indices. Applied Research in Environmental Economics, ZEW Economic Studies, 31, 2-22.

(2011.8.23受理)

#### Comparability of the Indicators in Comparative Performance Measurement

Tatsuya Ono

Faculty of Regional Studies, University of Tottori onotatsu@rstu.jp

#### **Abstract**

Local government benchmarking, or performance-comparison projects in local government will play an important role in the local government management of Japan. Since local government benchmarking and social indicators that compare regions have many methodological points in common, they are called comparative performance measurement (CPM) in combination. In CPM efforts, comparability of the indicators used for comparison is a matter of crucial importance. This article classifies the method of comparison in CPM and discusses the problems and applicability of various methods of standardization and adjustment from the viewpoint of comparability in comparing organizational performance. Special attention is paid to the scale types of measures. This article also reports findings from a study on the indicators generated by various CPM projects.

#### **Keywords**

comparative performance measurement, comparability, scale types, standardization, adjusted performance measures

#### 【研究論文】

### 都市行政評価ネットワーク会議 自治体業務改善のためのベンチマーキング

熊倉 浩靖 群馬県立女子大学 npo-kuma@xp.wind.jp

#### 要約

基礎自治体行政評価の行き詰まりを突破しようと、15の都市から始まった都市行政評価ネットワーク会議も7年目を迎え参加自治体は80以上となり、20の施策群を対象に偏差値と自治体別カルテの導入ができるまでに成長した。基本指標(当該事業を取り巻く状況)、結果指標(自治体としての施策業務内容=行政サービスの投入とその結果)、成果指標(行政業務に対する住民の対応=利活用度等)、コスト指標という4種類の指標群からなる構成に基づく比較評価手法は、自治体業務改善のツールとしての有効性が確認されつつある。

成人健康診査事業、介護保険事業、図書館サービス、文化会館管理・運営事業、消防・救急サービス、 ごみの収集・処分サービス、男女共同参画社会づくりを例として、会議が開発したモデル指標群の概要を 提示すると共に、最終アウトプットである自治体別カルテ表現方法の試行錯誤と活用例を示し、ベンチマ ーキングによる自治体業務改善の成果と課題をとりまとめた。

#### キーワード

基礎自治体行政評価(施策評価)、自治体間ベンチマーキング、 自治体の自発的参加、基本・結果・成果・コストの4構成、自治体別カルテ

#### 1. 基礎自治体行政評価の課題と解決方向

#### (1) 総務省調査が示唆する課題

2010年10月1日現在の総務省調査によれば、大阪府を除く46都道府県、さいたま市を除く18指定都市、函館市・高知市を除く38中核市、41の全特例市で行政評価が導入されているのに対し、一般の市区町村における行政評価の導入は、709市区のうち554で78.1%、941町村のうち280で29.8%となっている」。

さらに行政評価導入の内実に踏み込むと、事務 事業評価は、都道府県、指定都市、中核市、特例 市、市区、町村を問わず、ほぼ全ての自治体で行 われているが、施策評価は、都道府県89.1%、指 定都市61.1%、中核市63.2%、特例市56.1%、市区 46.9%、町村30.7%と、規模が小さくなるにつれ て逓減している<sup>2</sup>。

この数値だけで一般の市区町村の行政評価導入 が遅れているとは言い切れないが、住民に最も近 い所での行政サービスの責を負っている一般市区 町村にとって行政評価の導入とそれに基づく行政 改革推進にはなお大きな課題が横たわっていると 言わざるをえない。

その原因は何か。総務省調査では未導入自治体から回答を得ていないので推測となるが、自治体の規模に関係なく、導入済み自治体が掲げる「行政評価の課題」のトップが「評価指標の設定」となっていることは示唆深い。導入済み自治体に占める割合は、都道府県78.3%、指定都市100.0%、中核市84.2%、特例市85.4%、市区80.7%、町村73.2%である。

アンケート等の分析ではなく都市行政評価ネットワーク会議内での意見ではあるが、市区町村独自の指標設定がほとんどで、指標設定の適正さや指標値・達成度の全国位置を確認できないため、施策改善や総合計画の進行管理に生かせない状態におかれ、評価疲れさえ生じていると担当者は語っている。特に、法制度上、基礎自治体が担わざるをえない施策群において、その傾向が強い。

#### (2) 解決策としてのベンチマーキング

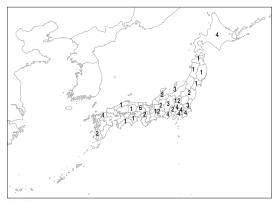
言い方を変えれば、各自治体では、施策、事務 事業に即した形で行政評価システムを導入したも のの、①各市が独自に評価方法を採用しているた め、他の自治体との比較・検討が難しく、いわゆ るベスト・プラクティスという行政評価の特性を 生かしにくいこと、②住民に身近なサービスに対 する評価例が少なく、住民にとって親しみにくい 評価体系になりがちなため、住民の参加・協働が 得にくいなどの問題が発生しているも。

あるいは「評価」という言葉が持つ価値付けと 行財政改革=業務・コスト縮減の意識から、業務 を廃止・縮減するためだけの行政評価イベントが 独り歩きし、持続的に実施すべき業務の効果・効 率的な施策改善には役立っていないという現象も 出始めている。

その解決策として、評価の導入目的の明確化と目的に見合った制度設計、既存の評価制度の改善・活用と並んで、基礎自治体間でのベンチマーキングが求められ出した<sup>5</sup>。

自治体間ベンチマーキングとは「他自治体との 比較により、評価の活用に資する知見を得る手法」 である。そこで私たちは、2001年度の総合研究開

#### 図1 ネットワーク会議参加自治体



(出所)筆者作成

発機構の研究を嚆矢に、米国ICMA(International City/County Management Association 国際都市経営協会)の業績測定比較(comparative performance measurement)モデルなどを参考に、コスト情報を重視したベンチマーク・モデルを提案・導入し、2005年11月15市を発起人に都市行政評価ネットワーク会議(Municipal Evaluation Consortium)を発足させた。12の施策から始めたが、参加市区町は80を超え対象施策も20に広がっている(図1)。指定都市・中核市・特例市を含む市区中心なので、80という数は全市区809の1割に当り、数値が示す基準性や偏りの評価に妥当性が得られる水準に近付いている。

筆者は、この過程で分析チームのリーダーとなり、データの集計・分析、後述する自治体カルテの作成に携わってきた。

ベンチマーキングの対象を施策評価としたのには、2つの理由がある。第1は、事務事業では数が多すぎ、かつ自治体の捉え方の差異が大きく比較しにくいこと、第2は、根拠法令等に基づいての全国比較が容易で、総合計画の進行管理や住民説明は概ね施策レベルで行われているからである。

#### 2. ベンチマーク・モデルの概要と特色

#### (1) 20のベンチマーク・モデル群

自治体業務のバランスを考えて、自治体が法令 上持続的に実施せざるをえない業務を対象に、社

		***
	施策	測定内容
社	成人健康健診事業	特定健診・がん検診等
	妊婦·乳幼児健診事業	法定検診·予防接種等
会	市町村国保事業	国保特別会計全体を比較
会保障系	保育所サービス	市立・民間比較等
系	学童保育サービス	設置·運営形態比較等
	介護保険事業	介護保険特別会計全体を比較
	スポーツ施設管理運営事業	なんなおうない 利田老安させ田
社会	図書館サービス	社会教育系は、利用者率を成果 指標とし、都市間比較と合わせ て、都市内部の各施設運営状況 を比較できる形としている
社会教育系	公民館活動	
育系	文化会館管理・運営事業	
,,,	博物館等管理・運営事業	
	市町村道維持・管理事業	交通安全面に評価の重点
基般	消防・救急サービス	消防・救急総体を比較
整	水道施設整備・運営事業	事業ガイドライン準拠
盤整備系	汚水処理整備・運営事業	下水等ガイドライン準拠
//	ごみ収集・処分サービス	ごみ処理事業総体を比較
参加·協働系	情報公開事業	広報事業評価に重点
	届出・証明交付サービス	戸籍関係証明で窓口比較
	男女共同参画社会づくり	男女雇用機会均等を評価
	国際化対応	国際理解等

表1 採用のベンチマーク・モデル群

(出所) 筆者作成

会保障(民生・福祉)系6施策、社会教育系5施策、 基盤整備系5施策、参加・協働系4施策を選び20の ベンチマーク・モデル群を作成した(表1)。

その際、多くの自治体が加工せずに比較的簡単に出せる公開データを基本に据え、議会認定の決算書をベースとした。コストデータが含まれていることと、誰もが同じ指標値に至れることが重要と考えたからである。

## (2) 4種類の指標群からなる構成

施策ごとのベンチマーク・モデルは、基本・結果・成果・コストの4種類の指標群からなる構成となっている(表2)。

表2 4種類の指標群からなる構成

指標	内容
基本	当該事業を取り巻く状況→自治体の特性、資産、課題等を
奉平	浮き上がらせる。
	自治体としての施策業務内容(行政サービスの投入とその
結果	結果)。企業経営で使われるinputとoutputが共に出てくる場
	合もある。
	行政業務に対する住民の対応(利活用度等)。企業経営で求
成果	められるoutcomeではなく、あくまでも住民の利活用度を測
	定、比較。
コスト	自治体としての施策業務内容と行政業務に対する住民の対
コスト	応をコスト面で比較。

(出所) 筆者作成

まず「当該事業を取り巻く状況」を基本指標とし、成人健康診査なら高齢化率や特定健診対象者率、水道事業なら水源構成を捉えて当該自治体の特性、資産、課題等を可能な限り計数的に浮かび上がらせることとした。

次いで「自治体としての施策業務内容(行政サービスの投入とその結果)」を結果指標とし、自治体としてどれだけの業務を行っているかを比較することとした。その結果、企業経営等で使われるinputとoutputが共にここから出てくるケースも見られることとなった。

他方、成果指標としては「行政業務に対する住民の対応(利活用度等)」を位置づけ、企業経営等で言われるoutcomeではなく、あくまでも住民の利活用度を測定、比較した。

そして「自治体としての施策業務内容と行政業務に対する住民の対応をコスト面で比較」するものとしてコスト指標を位置づけた。

この構成は、企業経営で用いられるベンチマークのinput・output・outcomeとは明らかに異なる構成となっており、企業経営のinput・output・outcome構成を踏襲しようとした多くの自治体の事務事業評価の指標構成とも異なる形になっている。自治体行政評価の指標設定の難題の一つに結果指標と成果指標の位置づけがあるが、その原因は、企業経営でのinput・output・outcome構成をそのまま当てはめようとしたためと考えられたからである。そこでinput・output・outcomeにこだわらず、自治体業務の結果は自治体側に、成果は住民側にと、視点を明確にすることで、自治体施策が適量かどうか、効果・効率的かを両者の比較でみる形とした。結論的に言うと、この仕掛けは成功している。

それらをコスト面で把握する指標がコスト指標で、効果的でも高コストの場合どう考えるか、各市が出し合った平均値や偏差値から多面的に見ていこうというわけである。

## (3) 集計・分析で初めて出てくる基準値

分析は、各市データを受けて平均や最大値、最小値、度数分布を求めることから出発する。初めから基準値が設定されているわけではないからである。当初は定義の共有が充分でなかったため大

34 能倉 浩靖

きなばらつきが見られたが、現在では偏差値や有 意性の高い度数分布が得られる指標が増えている。

全体分析の年次比較によって、基礎自治体全体 としての経年変化が望ましい方向に進んでいるか どうかも分かってきた。従来は、ある自治体が望 ましい経年変化を示したとしても、それが全体の 変化の枠内なのか、枠を超えたベターな変化なの か分からなかった。それが見えるようになった。

#### 3. 20のモデル群の概観

20のモデル群について、4分野ごとに1~2ずつ を詳述し他は概要を述べることとする。

#### (1) 社会保障系

社会保障系では、成人健康診査事業と介護保険 事業を例として取り上げたい。

成人健康診査事業は特定健診とがん検診を主対 象として表3の指標群を比較する。構成的には、 高齢化率と特定健診対象者率の高さを見ることで 当該自治体の特性を把握し(基本指標)、その状 況に対して実施されている特定健診やがん検診の 程度を受診者率で見 (結果指標)、その結果、特 定保健指導はどの程度行われ効果が出ているか、 死亡率や死亡原因構成比に成果が出ているかを検

表3 成人健康診査事業指標モデル

基	高齢化率
本	全人口に占める特定健康診査対象者の割合
	全人口に占める特定健康診査受診者の割合
	特定健康診査実施率(受診者数/対象者数)
結	受診者に対する特定保健指導対象者の割合
果指	特定保健指導対象者の構成比 (動機づけ・積極)
標	全人口に占める各種がん検診の延べ受診者率
	乳・胃・子宮・大腸以外に実施のがん検診
	全人口に占めるその他の成人健診受診者率
	特定保健指導実施率 (被指導者数/対象者数)
成	特定保健指導実施構成比(動機づけ・積極)
果指	10万人対死亡率
標	3大死因+糖尿病死亡率
	3大死因+糖尿病死亡者の構成比
	特定健康診査受診者1人あたりコスト
コス	特定健康診査受診者負担割合
1	各種がん検診受診者1人あたりコスト
指標	各種がん検診受診者1人あたり負担額
,,,,	成人健康診査総受診者1人あたりコスト

(出所) 筆者作成

討する (成果指標)。そしてコストを比較する。

そのために表4のデータ提供をいただく。表4は 項目が多いが、特定健康診査とそれに基づく特定 保健指導の対象者・受診者・被指導者数、各種が ん検診の受診者数、死亡者数とその構成(特定健 診の指導対象である悪性新生物・心疾患・脳血管 疾患・糖尿病)、特定健診・特定保健指導の総事 業費・総負担額、がん検診の総事業費・総負担額 などで各自治体が手元に持っている数値である (死亡原因統計は国からの還付数値なので遅れる)。 そこから表3を作っていく。各市データを1枚の

表4 成人健康診査事業入力表

人口 (年度末)	住民基本台帳+外国人登録
7 (1/2/19)	65歳以上人口
特定健康診査	対象者数
11 ACREMENT EL	受診者数
	動機付け支援対象者数
	積極的支援対象者数
特定保健指導	支援対象者総数 (自動計算)
刊是体展拍导	動機付け支援実施者数
	積極的支援実施者数
	指導実施者総数 (自動計算)
	胃がん
	子宮がん
	乳がん
タ番ぶり 炒シムシンサ	大腸がん
各種がん検診受診者数	前立腺がん
	肺がん
	その他のがん検診
	延べ人数 (自動計算)
	歯周疾患
その他の成人健診受診	骨量
者数	その他のその他成人健診
	延べ人数 (自動計算)
	全死亡者数
	悪性新生物による死亡者数
死亡者数	心疾患による死亡者数
	脳血管疾患による死亡者数
	糖尿病による死亡者数
特定健診総事業費等	総事業費 (関係職員人件費含む)
(千円)	受診者・被指導者負担総額
がん検診総事業費等	総事業費 (関係職員人件費含む)
(千円)	受診者負担総額
その他健診事業費等	総事業費 (関係職員人件費含む)
(千円)	受診者負担総額
成人健診総事業費等	総事業費(自動計算)
(千円)	総負担額(自動計算)
() 31.) # 1.#[24:31.0]	

(注記1) 死亡原因統計は暦年で、かつ1年ほど経ってから国から確定 値が示されるので、前年数値でも可。

(注記2) 健診事業費の内訳に関し按分等が難しい場合は「成人健診総 事業費」に直接入力し内訳項目は空欄。

excel表<sup>7</sup>に写し取り各指標値や指標値ごとの平均 値等を計算する。問題は偏差値が導入できるか、 つまり正規分布になっているかで、これは専門ソ フトに依拠している<sup>8</sup>。

各自治体には、後述するカルテに平均値等と合わせて各市の数値、偏差値を示し、11段階。に分けた指標ごとの評価、経年変化、4指標を通観しての改善点や課題、改善のために参考となる他市事例をコメントして自治体に返す。ベスト・プラクティスに学ぶガイドだが、結構ハードな仕事である。

成人健診事業は定義が明確なこともあって、ほ は全ての項目で偏差値が導入できており、各自治 体は、当該自治体の状況が標準的なのか、進んで いるのか、改善課題があるのかを見出せるまでに なっている。とくに特定保健指導等において、ど の成人病対策に当該自治体としては力を加えてい くべきかが見え始めてきたことは大きな成果と考 えられる。

特定健診は08年度から始まったばかりだが、国が示している2012年度実施率参考目標値60.4%<sup>10</sup>に対して、各自治体提供の09年度データでは最大でも56.7%、平均では35.9%であった。国目標に従うなら、さらなる普及・啓発が不可欠だが、当該自治体だけが低いのではなく全国共通の課題であることが確認できた意味は大きい。ベンチマークという手法を導入したからこその成果である。

類似の構成となっている**妊婦・乳幼児健康診査** 事業も多くの項目で偏差値が導入できており、周 産期死亡率・乳幼児死亡率と受診者率の変化など を対比できるまでになっている。

介護保険事業は、表5が提供いただくデータ、 表6が指標群となる。

構成としては、高齢化率で当該自治体の基本特性を把握し、要介護・要支援認定者の状態と介護事業従事職員体制を結果指標とした。職員配置は行政業務において自由度の高いリソースだからであり、単に人件費が高い、低いではなく、どのような質と量をもった職員を配置しているかを見ることに重点を置いている。他のモデル群を含めて、この設定は行政改革への繋がりを意識している。

他方、実際にどれだけのサービスを受けている かを成果指標とし、在宅・地域密着・施設に分け、

#### 表5 介護保険事業入力表

人口(年度末)	住民基本台帳+外国人登録
711-11-21-17	65歳以上人口
要介護・要支援認定者数	女 (報告値)
在宅介護者数(報告値)	
地域密着型介護者数(執	B告値)
施設(入所)介護者数	(報告値)
在宅(居宅)介護サーヒ	ごス給付金 (千円)
地域密着型介護サービス	(給付金(千円)
施設介護サービス給付金	全 (千円)
	通所介護 (デイサービス)
	通所リハビリテーション (デイケア)
在宅介護施設定員	短期入所生活介護
	短期入所療養介護
	総数 (自動計算)
地域密着型介護施設	特定施設入所者生活介護
定員	認知症対応型共同生活介護
上月	総数 (自動計算)
	介護老人福祉施設(特養ホーム)
   施設介護施設定員	介護老人保健施設
旭叹月夜旭叹疋貝	介護療養型医療施設
	総数 (自動計算)
歳入予算	総額 (千円)
NA   异	保険料見込み額 (千円)
歳入決算	総額 (千円)
/////////////////////////////////////	保険料収入済額 (千円)
一般会計からの繰入金	(千円)
	正規職員数
従事職員数	嘱託・臨時・再任用等
	総数(自動計算)

(注記) 要介護・要支援認定者数等は年度末月数値で可

(出所) 筆者作成

#### 表6 介護保険事業指標モデル

基本	高齢化率
結	全人口に占める要介護・要支援認定者の率
果指	従事職員1人あたり要介護・要支援認定者数
標	従事職員の構成(正規・嘱託等)
	在宅介護率(要介護・要支援認定者対比)
	在宅介護施設定員に対する在宅介護者比率
	在宅介護者1人あたり在宅介護給付金
成	地域密着型介護率(要介護・要支援認定者対比)
果指	地域密着型施設定員に対する地域密着型介護者比率
標	地域密着型介護者1人あたり地域密着型介護給付金
	施設介護率(要介護・要支援認定者対比)
	施設介護施設定員に対する施設介護者比率
	施設介護者1人あたり施設介護給付金
	要介護・要支援者1人あたり換算保険料収入
_ [	保険料徴収率(収入済額/見込み額)
コス	予算額に対する決算額の割合
上	歳入に占める保険料の割合
指標	要介護・要支援者1人あたり決算額
	要介護・要支援者1人あたり一般会計からの繰入額
	一般会計からの繰入率

それぞれに介護者率、施設定員に対する被介護者の関係、被介護者1人あたりサービス給付金を比較している。この比較は、他自治体との間だけでなく、当該自治体内において住民需要とサービス供給とが合致しているかを見る目安ともなっており、社会教育系の一連のモデル群でも援用している。

本モデル群でも偏差値が導入できる指標が増えているが、介護専担部局では常識となっていると思われる多くの数値を、自治体職員全体、議員、住民間で共有する役割も果たすこととなった。

因みに平均値で言えば、要介護・要支援認定者の6割前後が在宅介護者で、在宅介護施設定員の3倍程度の介護需要があり、在宅介護者1人あたりの給付金は110万円程度である。同じく5%が地域密着型介護者で、地域密着型施設定員の1.5倍の需要、1人あたり給付金は230万円程度、施設介護者は2割弱で、数値上では施設介護定員と需要は合致、給付金は310万円程度である。

コスト指標は決算額、保険料収入や徴収率の実態に合わせ、一般会計からの繰入率・繰入額を重視した。介護保険特別会計としての持続的自立性担保を意識してだが、平均値で、国が費用負担として定めている12.5% で若干上回る14%という数値が出ている。

市町村国保事業モデルは類似の構成を取り、こちらの繰入率は平均値で7%程度である。原単位での決算額、歳入に占める保険料割合、一般会計からの繰入率の各偏差値を比較・監視しベスト・プラクティスに学んでいくことは、両特別会計の持続的自立性を図るため極めて有効な方法と見られ、ベンチマーキングが本領を発揮する場面と言ってよい。

社会保障系の他の2つのモデル群は保育所サービスと学童保育サービスで、それぞれ、公立と民間、公設公営・公設民営・民設民営に分けて、サービス提供量や住民利用の実態、児童1人あたりコストなどを比較する構成としている。全国的な傾向を言えば、保育所サービスでは民間が、学童保育サービスでは公設民営タイプが、サービス提供面でも住民利用の実態でもコストでも比較的優れた数値を出し続けている。市場化テストの話が持ち出されるが、保育所サービス、学童保育サービスに関しては、あるべき施設設置・運営形態は

ほぼ見えてきていると言えよう。

#### (2) 社会教育系

社会教育系の5つのモデル(スポーツ施設管理 運営事業、図書館サービス、公民館活動、文化会 館管理運営事業、博物館等管理運営事業)は類似 の構成を取る。

図書館サービスと文化会館管理運営事業を例として説明したい。図書館サービスの入力表は表7、 指標モデル群は表8、文化会館管理運営事業は表9 と10である。

両施策の指標モデル構成は同じで、施設の数や1館あたり平均サービス人口、指定管理者導入の状況を基本指標、自治体におけるサービス供給の前提とし、職員配置と、図書館なら住民1人あたり蔵書・ソフト数、文化会館なら住民10,000人あたり座席数(会議室等含む)などの施設固有のサービス供給量を結果指標、住民の利用状況を成果指標とした。

図書館サービスで利用者(入館者)率と貸出者率とを別立てしたのは、閲覧・リファレンスサービスなどの図書館の機能拡張、総合情報館としての役割を把握したかったからである。

文化会館管理運営事業では座席回転率という概念を導入した。稼働率ということがよく言われるが、自治体によって定義がばらばらで、かつ稼働

表7 図書館サービス入力表

総人口 (年度末)	住民基本台帳+外国人登録
図書館数	
指定管理者導入の有無	
図書館協議会設置の有無	Ħ.
蔵書・ソフト数	年度末実数
殿音・ ノノト奴	当該年度増加数
	正規職員数
	嘱託・臨時・再任用等
配置職員数	指定管理者職員
	総数 (自動計算)
	うち司書有資格者数
施設年間利用者数(入場	易者総数)
図書等貸出者総数	
年間図書等貸出総数	
管理運営総事業費(人作	‡費を含み減価償却費は含まず)
年間図書等購入費	
( \\ \_ \_ \)	一 1 が 曲 4 人 3 体 体 が 曲 4 人 4 4 、 1 1

(注記) 管理運営総事業費で人件費を含み減価償却費を含まないとしたのは運営を見るのが目的だからであり、同様の観点から、 建設費や大型の補修費は除外していただきたい。

表8 図書館サービス指標モデル

-11-	図書館数
基本指標	図書館協議会設置の有無
指標	指定管理者導入の有無
ীয়াং	1館あたり平均サービス人口
	配置職員1人あたり平均サービス人口
<b>♦±:</b>	職員構成比(正規・嘱託等・指定管理者)
結果指標	1館あたり配置職員数
指標	配置職員に占める司書有資格者比率
1নং	住民1人あたり蔵書・ソフト数
	住民10,000人あたり年間蔵書等増加数
成	全人口に対する利用者(入館者)率
果	全人口に対する貸出者率
果指標	全利用者(入館者)に占める貸出者の割合
ানং	住民1人あたり年間図書等貸出数
コ	貸出者1人あたり管理運営コスト
スト	利用者(入館者)1人あたり管理運営コスト
指	住民1人あたり管理運営コスト
標	住民1人あたり図書等購入費

(出所) 筆者作成

#### 表9 文化会館管理・運営事業入力表

総人口 (年度末)	住民基本台帳+外国人登録
文化会館 (類似施設) 娄	¢
指定管理者制度導入施設	<b>发数</b>
総座席数	会議室含む
	正規職員
新1 黑 聯. 吕 米/s	嘱託・臨時・再任用等
配置職員数	指定管理者職員
	総数(自動計算)
技術専門職数	舞台・音響・照明の専門職数
	自治体·文化会館主催事業
年間入館者数	住民団体主催事業
中间入路有效	その他
	総数(自動計算)
管理運営総事業費(人作	-   費・委託費は含み減価償却費は含まない)
利用者からの収入総額	(物販販売等も含む)

(出所) 筆者作成

表10 文化会館管理・運営事業指標モデル

基本指標	文化会館数
指標	指定管理者導入率
	住民10,000人あたり座席数
結	1館あたり平均サービス人口
果指標	1館あたり配置職員数
標	職員構成比(正規・嘱託等・指定管理)
	職員に占める舞台・音響・照明等専門職の率
成	全人口に対する利用者率
果指標	入館者構成 (主催・住民団体・その他)
標	座席回転率
コ	利用者1人あたり管理運営コスト
スト	住民1人あたり管理運営コスト
指	利用者1人あたり支出額
標	管理運営費に対する利用者支出額割合

(出所) 筆者作成

率という計算された数値が提出され、同一の定義 に基づいた実際の利用実態が把握しにくいためで ある。飲食業の座席回転率を援用したが、比較の 上で効果を発揮している。

コスト評価では、住民1人あたりと利用者1人あたりとを別立てして比較する方法を採った。前者、いわば投資コストが比較的高めでも、後者、いわば利用コストが低ければ、効果・効率的な運用となっている訳で、運用の実態を的確に示すコスト指標となった。

この構成と分析の方法は、社会教育系の5つのモデルに共通しており、そこから、他自治体との比較だけでなく、自治体内における各施設の運用・利用の比較もできるようになっている。ただし現状では、博物館等と公民館のモデルの熟度はまだ高くない。

平均値での比較ではあるが、5つのモデルを比較すると、利用者率では、スポーツ施設が440%(住民は1年間に4.4回スポーツ施設を利用するということで、以下同様)で一番高く、図書館290%、公民館240%、文化会館160%、博物館等80%と続く。住民は、博物館等へ行く5倍以上の頻度でスポーツ施設を利用し、図書館は4倍弱、公民館は3倍、文化会館は倍利用するのが一般傾向ということが分かったことになる。

指定管理者の導入は文化会館とスポーツ施設で高く、公民館、図書館、博物館等では低い。3施設が低いのは、設置法令が明確で、機能の純化や専門職の配置が規定されており、かつ利用者負担も低いためであろう。こうしたことが分かったのもベンチマークを導入したからで、施策改善議論の目安となろう。

#### (3) 基盤整備系

基盤整備系は、市町村道維持・管理事業、消防・救急サービス、水道施設管理運営事業、汚水処理施設管理運営事業、ごみ収集・処分サービスとあるが、モデルとして熟度の高い消防・救急サービスとごみ収集・処分サービスを例に挙げる。

表11が消防・救急サービスの入力表、表12が指標モデルである。消防・救急事業、消防署・消防団全体を対象とした。全体で住民の安全が守られているからである。

38 熊倉 浩靖

表11 消防・救急サービス入力表

総人口 (年度末)	住民基本台帳+外国人登録
面積	km <sup>2</sup>
消防署数	常備消防
消防団分団数	非常備住民組織
消防署員数	年度末実数
救急救命士数	年度末実数
消防団員数	年度末実数
稼働している消防車数	消防署配備
(指令車除く、年度末	消防団配備
平均台数)	総数 (自動計算)
稼働している救急車数	(年度末平均台数)
高規格救急自動車数	
年間火災発生件数	消防車出場件数
年間焼失面積 (m²)	全面積
平间烷大凹槓 (m)	林野面積
火災損害額 (千円)	全額
八次損害額(十円)	林野損害額
救急車出場件数	実数
救急車到着までの所要年	F間総時間 (分)
救急車到着から搬送完丁	「までの所要年間総時間 (分)
常備消防総事業費	減価償却、施設の新・改築、大規模改
消防団総事業費	修は含まない。
	· · · · · · · · · · · · · · · · · · ·

(出所) 筆者作成

表12 消防・救急サービス指標モデル

基	消防署1署あたり平均サービス人口	
本	消防署1署あたり平均サービス面積	
指標	消防団1分団あたり平均サービス人口	
际	消防団1分団あたり平均サービス面積	
	消防署員1人あたり平均サービス人口	
44-	消防署員に占める救急救命士比率	
結果	消防団員1人あたり平均サービス人口	
指標 稼働している消防車1台あたり平均人口		
际	稼働している救急車1台あたり平均人口	
	救急車に占める高規格救急自動車の割合	
	住民10,000人あたり火災発生件数	
-4-	住民10,000人あたり林野を除く焼失面積	
成果	住民10,000人あたり林野を除く火災損害額	
指標	住民10,000人あたり救急車出場件数	
救急車平均到着時間		
	救急車平均搬送時間	
コスト	住民1人あたり常備消防費	
指標	住民1人あたり消防団経費(非常備消防費)	

(出所) 筆者作成

指標モデルの構成は、消防署・消防団の配置を 平均サービス人口と面積の両面から見ることを基本指標とし、消防職員・消防団員・消防車・救急 車の配置を結果指標、火災発生件数・焼失面積・ 火災損害額・救急車出場件数と救急車平均到着・ 搬送時間を成果指標とし、住民1人あたり経費で コスト比較とした。 消防・救急サービスで注目されるのは救急車の 平均到着時間である。10,000人あたり出場件数は ほぼ横ばいなのに対し、到着時間は年々遅くなっ ている (表13) <sup>12</sup>。6年間で1分も遅くなっている ことの原因解明が必要と見られるが、こうした全 体的な問題点が見えてきたことも、ベンチマーキ ングの成果と言えよう。

表13 救急車出場平均値の経年変化

年度	04	05	06	07	08	09
出場件数	370.0	376.5	369.5	375.4	356.3	364.4
到着(分)	6.2	6.3	6.6	6.7	6.9	7.2

(出所) 筆者作成

ごみの収集・処分サービスでは、表14が入力表、 表15が指標モデルとなっている。

指標モデルの構成は、ごみ処理の流れに即して、 地域で出されているごみの量を基本指標、焼却・ 処分場の状況を結果指標、リサイクルと最終処分 の実態を成果指標とし、コスト指標で経費と手数 料等を比較した。

環境省に報告している一般廃棄物処理実態調査数値があるためか定義の揺らぎが少なく、ほとんどの指標で偏差値が得られているが、ごみ排出量は平均値で年々大きく減少し、国民・企業のごみ減量の意識と実践が確実に進んでいることが確認できる(表16)<sup>13</sup>。逆に言うと、各自治体でごみ減量が進んだと見ても、全国平均を下回っている

表14 ごみの収集・処分サービス入力表

総人口(年度末)	住民基本台帳+外国人登録
分別種類	環境省一廃処理実態調査基準種別準
	生活系ごみ排出量
収集量	事業系ごみ排出量
	総量 (自動計算)
総排出量	総収集+集団回収(自動計算)
直接焼却量	報告値
焼却施設の種類	焼却、ガス化溶融・改変、なし
焼却施設余熱利用	温水・蒸気・発電・その他
	直接資源化量
11 11 / 4 4 日	中間処理後再生量
リサイクル量	集団回収量
	総量(自動計算)
最終処分量	報告値
最終処分場	処分場数
<b>公市労典 (1円)</b>	最長残余年数 (年単位)
総事業費 (千円)	運営・維持のみ。建設費含まず。
手数料等収入総額(年間	引、千円)
(.t. ⇒.e) to be to b	

表15 ごみの収集・処分サービス指標モデル

住民1人1日ごみ辨出量 (総・株本学) 住民1人1日ごみ辨出量 (総・排出) 住民1人1日ごみ焼却量 若		分別種類					
住民1人1日ごみ辨出量 (総・株本学) 住民1人1日ごみ辨出量 (総・排出) 住民1人1日ごみ焼却量 若	基	住民1人1日ごみ排出量(生活系)					
住民1人1日ごみ辨出量 (総・株本学) 住民1人1日ごみ辨出量 (総・排出) 住民1人1日ごみ焼却量 若	半	住民1人1日ごみ排出量(事業系)					
住民1人1日ごみ焼却量  お	標	住民1人1日ごみ排出量(総収集)					
お こみ焼却施設の種類(焼却・ガス化・なし) 焼却施設余熱利用(温水・蒸気・発電・その他) 最終処分場の数 最終処分場最長残余年数 住民[人]日リサイクル量 リサイクル率 リサイクルの構成(直接・中間・集団) 住民[人]日最終処分量(直接埋立除く) コス 住民[人あたりごみ処理総コスト		住民1人1日ごみ排出量(総排出)					
標 最終処分場の数 最終処分場最長残余年数 住民1人1日リサイクル量 リサイクル率 リサイクルの構成 (直接・中間・集団) 住民1人1日最終処分量 (直接埋立除く) ユス 住民1人あたりごみ処理総コスト		住民1人1日ごみ焼却量					
標 最終処分場の数 最終処分場最長残余年数 住民1人1日リサイクル量 リサイクル率 リサイクルの構成 (直接・中間・集団) 住民1人1日最終処分量 (直接埋立除く) ユス 住民1人あたりごみ処理総コスト	結	ごみ焼却施設の種類 (焼却・ガス化・なし)					
標 最終処分場の数 最終処分場最長残余年数 住民1人1日リサイクル量 リサイクル率 リサイクルの構成 (直接・中間・集団) 住民1人1日最終処分量 (直接埋立除く) ユス 住民1人あたりごみ処理総コスト	米   指	焼却施設余熱利用(温水・蒸気・発電・その他)					
成果	標	最終処分場の数					
成 果 ガサイクル率 ガサイクルの構成(直接・中間・集団) 住民1人1日最終処分量(直接埋立除く) コス 住民1人あたりごみ処理総コスト		最終処分場最長残余年数					
果 リサイクル率 指標 リサイクルの構成 (直接・中間・集団) 住民1人1日最終処分量 (直接埋立除く) コス 住民1人あたりごみ処理総コスト	rt;	住民1人1日リサイクル量					
住民1人1日最終処分量(直接埋立除く) コス 住民1人あたりごみ処理総コスト	果	リサイクル率					
住民1人1日最終処分量(直接埋立除く) コス 住民1人あたりごみ処理総コスト	指揮	リサイクルの構成(直接・中間・集団)					
□	1নং	住民1人1日最終処分量(直接埋立除く)					
住民1人あたりごみ処理手数料等   指   手数料等市民負担率	コフ	住民1人あたりごみ処理総コスト					
標 手数料等市民負担率	上	住民1人あたりごみ処理手数料等					
	標	手数料等市民負担率					

(出所) 筆者作成

場合はベスト・プラクティスに学ぶことが求められる。なお、数値としてのリサイクルは上げ止まりと見られ、また、数値による限り、ごみ減量が最終処分量の減少には直結しておらず、今後の検討課題となっている。

表16 ごみ排出量等全国平均値の経年変化

(住民1人1日あたりで、単位はg)

年度	04	05	06	07	08	09
生活系	727	719	696	687	674	655
事業系	335	304	300	275	261	253
総収集	1,062	1,022	996	962	935	907
総排出	1,142	1,092	1,072	1,028	1,001	971
焼却量	858	785	771	789	786	754
recycle	207	226	224	213	200	195
recycle率	18.5%	21.1%	21.1%	21.4%	20.4%	20.5%
処分量	191	152	97	105	204	97

(出所) 筆者作成

基盤整備系は、他に市町村道維持・管理事業、 水道施設管理・運営事業、汚水処理施設管理・運 営事業を指標モデルとしているが、数値根拠の定 義の共有等で熟していない部分があり修正を重ね ている。

## (4) 参加・協働系

参加・協働系として情報公開事業、届出・証明 交付サービス、男女共同参画社会づくり、国際化 対応の指標モデル化を図っているが、住民サービ スとしての設計が難しく熟度は低い。中で、新た な比較成果と波及効果が生まれた男女共同参画社 会づくりを例として取り上げたい。

男女共同参画社会づくりには制度整備、普及・ 啓発、相談・DV対策、活動団体支援など多面的 な側面があるが、まずは事業所としての自治体に おける男女雇用機会均等の現状評価を核とした指 標モデルとした。表17が入力表、表18が指標モデ ルである。

構成的には制度整備(条例等・推進計画等・担 当部署・相談窓口・審議会・DV防止基本計画・

#### 表17 男女共同参画社会づくり入力表

男女共同参画条例等の有無(条例、推進計画、担当部署、相談窓口、審議会、DV防止基本計画、DV相談支援センター)						
講座等、住民向け男女共同参画推進事業実施回数						
講座等、住民向け男女共同参画推進事業参加者数						
正規職員						
自治体職員数	嘱託·臨時·再任用等					
	総数 (自動計算)					
	正規職員					
女性職員数	嘱託・臨時・再任用等					
	総数 (自動計算)					
管理職数	係長担当職以上					
女性管理職数	係長担当職以上					
議会議員数						
女性議員数						
条例等で設置されている審議会等の委員数						
条例等で設置されている審議会等の女性委員数						
男女共同参画社会形成の	男女共同参画社会形成のための独自広報の有無					
男女共同参画社会形成に	係る団体支援の有無					
男女共同参画形成推進総	念事業費 (人件費含む)					

(出所) 筆者作成

表18 男女共同参画社会づくり指標モデル

	条例等の有無					
#	住民10,000人あたり自治体職員数(全体)					
基本指標	住民10,000人あたり自治体職員数(正規)					
	住民10,000人あたり自治体職員数(嘱託等)					
1分	住民10,000人あたり議員数					
	住民10,000人あたり審議会等委員数					
	女性職員率 (全体)					
	女性職員率 (正規)					
結果	女性職員率 (嘱託等)					
結果指標	住民10,000人あたり推進講座等開催数					
保	独自の広報の有無					
	独自の団体支援の有無					
	女性議員率					
成	女性管理職率					
成果指標	女性委員率					
標	住民10,000人あたり推進講座等参加者数					
	推進講座等1回あたり参加者数					
コスト指標	住民1人あたり男女共同参画事業コスト					
/.I \ &-6-	1. 1. 1.					

DV相談支援センター)と事業所としての自治体の概要を基本指標とし、女性職員率や推進講座等の開催数、独自の広報や団体支援の有無を結果指標、女性議員率・女性管理職率・女性委員率、推進講座等の参加者数を成果指標とした。職員把握で正規と嘱託等を別立てとしたのは、全体としては女性比率が低くはないが、内実、正規職員比率が低い自治体が散見されたからである。

こうした構成から男女共同参画社会づくり指標 モデルは、住民サービスという観点とは異なる比 較成果と波及効果を生んでいる。

まず制度整備では、担当部署や推進計画等はほとんどの自治体が作っており相談窓口や関連審議会も4分の3の自治体が作っているが、条例等の策定は4割に落ち、DV防止基本計画とDV相談支援センターを作っている自治体は1割に過ぎないことが分かった。決して望ましい状況とは言えない。

次に職員数は、平均値で、住民10,000人あたり 正規職員90人弱だが、嘱託等が50人弱あり、全体 では130人強あることが判明した。職員の4割ほど は非正規職員であること、行政改革における職員 数は、この全体像と内部構成を吟味して議論する ことが大切であることが分かった。一つの波及効 果と言える。

女性職員率では、正規職員の女性職員率は平均値で37.9%という数値がもたらされた。全体としては49.1%で、男女半々に見えるが、正規の女性職員は4割を割っているのが実情である。管理職率になると、さらに甚だしく、女性管理職率は18.8%である。係長担当職以上としており、女性職員比率の高い保育所や看護・介護職なども含まれていることを考えると、他部局での女性管理職率はかなり低い。

その一方で、女性委員率は、国の第2次・第3次男女共同参画基本計画が数値目標とした女性の国家公務員や国の審議会等の女性委員「2020年30%」はを基準に、平均値でも24.4%となっている。これ自体は望ましい動きと言えるが、事業所としての自治体の女性職員率とその構成や女性管理職率とのギャップを指摘せざるをえない。

なお、コスト指標も設けてあるが、現状、有意 な基準は見出せていない。

情報公開(公文書開示・広報等)、届出・証明

等窓口サービス、国際化対応はさらに熟度が低く、 参加・協働系指標は課題が多い。

## 4. 最終アウトプット:自治体別カルテ

このように、全体分析も示唆深いが、自治体が有益と感じているのは自治体別カルテである。20 の施策ごとに全体分析を踏まえた各自治体数値を位置づけ望ましい方向を明らかにし、「ベスト・プラクティスに学ぶ」の考え方から、可能であれば参考・目標となる他自治体を示すことがカルテの役割となっている。

## (1) カルテ表現方法の試行錯誤

各自治体の数値や全体の位置づけをどう表現するかについては試行錯誤を重ねた。一例を消防・ 救急サービス指標モデルの救急車平均到着時間で 示したい。

当初は、図2のようにヒストグラム上に各自治 体位置を直接に表現する方法を採った。

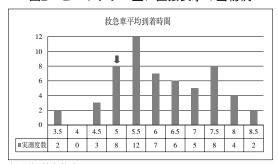


図2 ヒストグラム上に直接表示の当初例

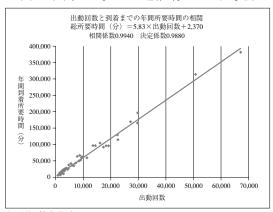
(出所) 筆者作成

矢印で示されている位置が、当該自治体の含まれている度数となる。直截で分かりやすいが、紙幅を費やし、他の数値との関連、全体像の把握が難しい。

参加自治体が増えるにつれ、0.95を超える高い 相関係数を示す理論直線が得られる例も出てきた (図3)。理論直線と当該自治体の位置を示すこと で、望ましい状況かどうかが判断できるケースが 出てきたことになる。

しかも経年変化が見やすく (表19)、理論直線

図3 相関から求めた理論直線上での表示例



(出所) 筆者作成

表19 理論直線による経年変化対比例

都下	<b></b>	相関係数	理論直線
03	03 15 0.9813		年間所要時間=5.23×出動回数+2,790
04	30	0.9841	年間所要時間=5.68×出動回数+1,780
05	57	0.9904	年間所要時間=5.83×出動回数+2,370
06	66	0.9538	年間所要時間=5.62×出動回数+3,316

(出所) 筆者作成

の傾きを見るだけでも傾向が分かる。

しかし、これもヒストグラムと同じ問題点を持 つと共に、理論直線が示せない指標との関係をど う見るかが課題となった。

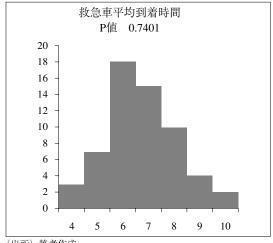
#### (2) 偏差値型の自治体カルテの導入

高い相関係数を示す理論直線が得られる、ある いは山型の度数分布であることは、正規分布と見 なせる可能性を示唆する。

そこで、有意水準5%での検定を行い帰無仮説 が採択できるかを、それぞれの指標で確認するこ ととした結果、幸いにして、多くの指標値で有意 水準5%の検定で正規分布とみなしてもよいとい う結果が得られたので偏差値導入となった(図 4・5は09年度データに基づく平均到着時間の有意 水準5%でのγ<sup>2</sup>検定と正規確率プロットの結果。 P値は0.7401の高さを示している)。

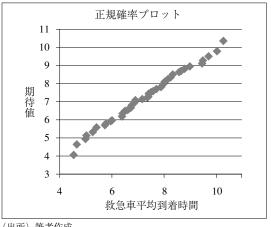
こうして各指標モデル1頁ずつ、計20頁の自治 体カルテができることなった。例としてT市の消 防・救急サービスを示した(図6)。

図4 有意水準5%での χ²検定例



(出所) 筆者作成

図5 図4データの正規確率プロット計算例



(出所) 筆者作成

#### (3) 現行自治体別カルテ

自治体別カルテは、それぞれの指標モデル群で、 評価の視点、当該自治体データの評価と全体の位 置、コメントの3つが書かれる。

評価の視点は、指標モデルの構成と変更事項、 項目評価の手法 (偏差値導入の有無)、全体的な 顕著な傾向を記すこととし、例えば消防・救急サ ービスでは、「消防隊員/救急隊員に分けること が難しいことから救急救命士比率は消防署員全体 に対する比とし、新たに火災損害額・焼失面積を 加えました。消防車・救急車あたりの人口、火災 損失額に関して新たに偏差値を導入することがで

42 熊倉 浩靖

#### 図6 偏差値型の自治体カルテ例

AH 高崎市様 09 年度カルテ

13 消防・救急事業 (消防・救急サービス:データ提供 64 市町)

■野毎の観点 前が探询や後急解判に分けることが確しいことから枚急後命士比申は前房署員会体に対する比とし、終亡に再写を徐く大災刑害前・焼失組積を加えました。 前房庫・穀卓生たりの人口、火災刑害第1に関して新たに延差触を爆入することができました。平均能レベルで、彼急郎出場合数はそれほど明えていませんが、到前時間、 施送時間とも 08年度に比べて延者相同が強まっていることが整念されます。

	町市データの評価と全体での位置						PR TRUE	100	MAN . IN THE
	抱標	08 年数值	御市歌値	個差值	評価	全国平均值	最大值	最小值	備考
	消防署1署あたり人口・面積	34,795 人	87,461 人			49,642 人	177,071 人	10,031 人	08 年度平均 47,371 人
基	消防署1署あたり平均サービス面積	40 kri	45.9 kd			110 kri	893 kri	4 krii	08 年度平均 100 kri
本	消防団 1 分団あたり平均サービス人口	6,959 人	6,042人			8,984 人	41,113 人	1,462 人	08 年度平均 8,498 人
	消防団 1 分団あたり平均サービス面積	/	7.4 Hi			14.0 km	63.3 kril	1.7 km	新規定義項目
	消防署員1人あたり平均サービス人口	1,074 人	1,061 人	54.6	+	942 人	1,542 人	390 人	外れ値1市、08年平均94.7人
	前防署員に占める教急教命士比率	11.7%	18.0%	44.0		17.0%	32.9%	0.0%	08 年度平均 15.8%
結	消防団員 1 人あたり平均サービス人口	290 人	274 人		標準域	292 人	1,208 人	26 人	標準城200~300人、08年297人
果	稼働している消防車1台あたり人口	4,519 人	4,257人	57.5	++	3,088 人	11,564 人	628 人	08 年度平均 2,961 人
	稼働している核急車1台あたり人口	26,765 人	26,758 人	50.4	標準的	26,195 人	59,957 人	7,300 人	08 年度平均 26,869 人
	教急車に占める高規格教急自動車割合	100.0%	100.0%			90.9%	回答 63 市中 1	00%46 市 73.0	%、08 年度平均 88.6%
	市民 10,000 人あたり火災発生件数	3.4件	8.8件			3.9件	8.1件	1.8件	標準城 3~4 件、08 年 4.2 件
	市民 10,000 人当り林野除く焼失面積	$\setminus$	152.9 ਜੀ			192 m²	1,693 m²	8 m²	新規定義項目
成	市民10,000人当り林野除く火災損害額		11,805,786円	59.5	++	7,217,885円	18,73,313円	763,698 円	新規定義項目、外れ値2市
果	市民 10,000 人あたり救急車出場件数	351.1 件	854.9 件	48.8	標準的	364.4 件	501.0 件	257.8件	08 年度平均 356.3 件
	救急車平均到着時間	6.1 分	7.4分	51.5	標準的	7.2 分	10.3 分	4.5 分	外れ値2市、08年平均6.9分
	救急車平均搬送時間	23.1 分	20.8分	41.6		25.1 分	39.9 分	16.0 分	外れ値 2 市、08 年平均 23.8 分
П	ス 市民1人あたり常備消防費	10,020 円	10,222 円	48.1	標準的	10,663 円	15,591 円	4,697 円	外れ値 4 市、08 年平均 11,118 円
1	市民1人あたり消防団経費	1,003円	1,240円	54.0	+	974円	2,807円	119円	外れ値2市、08年平均935円

■コメント 全体的に標準的で、救急車搬送時間が短いことは評価されますが、1 人あたり火災損失額が高めであること、救急救命土比比率が低めであることは改善課題と見られます。

都市行政評価ネットワーク会議

(出所) 筆者作成

きました。平均値レベルで、救急車出場件数はそれほど増えていませんが、到着時間、搬送時間とも08年度に比べて延着傾向が強まっていることが懸念されます。」と書いた。

当該自治体データの評価と全体の位置は、指標ごとに、前年度数値、当該年度数値、偏差値が導入できた場合は偏差値、先述した11段階の評価(注記9)、全国平均値、最大値、最小値、備考(前年度平均値や偏差値が導入できない場合に平均値・中央値・度数分布から判断した「標準域」を表示)を示すという形にしている。

指標ごとに全国での位置や経年変化が確認できるわけだが、全体としてどう評価すべきが課題となる。それがコメント欄であり、T市例では「全体的に標準的で、救急車搬送時間が短いことは評価されますが、1人あたり火災損失額が高めであること、救急救命士比比率が低めであることは改善課題と見られます。」と記した。特異な数値が出た場合は、原因等の可能性を示唆すること、可能な限り参考・目標となる自治体を例示することに努めているが、なかなか達成できていない。

#### (4) 自治体別カルテの活用例

作成して自治体に返した自治体カルテを当該自 治体の施策評価や総合計画進行管理の参考指標と する例が増えてきており、総合計画の施策指標の 一部に援用し出した自治体もある。その中で積極 的な活用例を2つ紹介したい。

第1は、カルテを使った庁内の研修で、私自身が呼ばれて説明することが多いが、関係部局あるいは全庁職員が集まりカルテを叩き台に数値や施策のあり方について討議する。

そのことによって、私自身が奇妙に思っていた数値の原因が判明したり、当該部局自体は問題・課題と思っていたが説明できず施策転換・財政要求ができないでいたことに計数的な裏付けができたりする例が増えている。一方で、課題は共有されながら解決策や学ぶべきベスト・プラクティス対象を示せない場合もあり、新たな課題が現場から生まれている。その検討を通してモデルのマイナーチェンジが年々行われるようになっている。

中には全庁横断型の若手職員研修として解決策 の提案を考えさせている自治体もあり期待が膨ら む。

第2は、関係部局に差し戻して当該部局にコメント(評価は正鵠を得ているか、実感等と合うなら課題設定と解決策をどう考えるか)を書かせ、全庁で情報交流を図っている例である。ここまで来ると、カルテは真に活かされていると実感する。

## 5. 都市行政評価ネットワーク会議の課題

しかし、このような例は希少である。ベンチマ

ーキングによる基礎自治体行政評価が進むために は、3つの課題が横たわる。

#### (1) 参加自治体増加による基準値の安定

第1は、参加自治体の増加による基準値の安定。 市区中心に80以上の参加があり市区の1割に達し ているとはいえ、市区町村1363に比べると5%程 度である。統計学的な安定をえるためにも、全体 の1割、130から150の市区町村の参加が得たい。

#### (2) モデル指標群の熟度向上とバランス

第2は、モデル指標群の熟度向上と基礎自治体施策体系と対応した指標群のバランス確保。例示したモデル指標群は熟度が高いものだが、参加・協働系などは全体として熟度が低い。また、2000年度以来何度となく挑戦しながら成案を得られず、どの分野においても産業政策系は全くモデル化できていないし、基盤整備系でも都市計画や公園管理等の漏れが多く、バランスが取れていない。

#### (3) 自治体間・関係部局間の意見交換

第3は、積極的な活用例の増加と自治体間・関係部局間での意見交換の拡大。例示した研修や当該部局コメントの導入などが進むと共に、まさに「ベスト・プラクティスに学ぶ」全国的な勉強会の広がりが期待される。現在も参加自治体の自発的参加によって全国規模での勉強会を開催しているが、担当部局レベルでの討議の深まりが期待される。

以上、都市行政評価ネットワーク会議の経緯、現状、成果、課題を書かせていただいた。非学会員であるにもかかわらず、2010年秋の第11回研究大会に発表の機会を与えられたばかりか、今回論文掲載を許されたことに感謝して擱筆としたい。

なお、本稿に示した図表は全て都市行政評価ネットワーク会議事業(ベンチマーク・モデルの設計、データの収集・分析、自治体別カルテ作成)の中で私が作成したものである。

#### 注記

1 平成23年3月16日付け総務省報道資料「地方公共団

体における行政評価の取組状況(平成22年10月1日 現在)」p1

- 2 上記総務省資料p4
- 3 上記総務省資料p7
- 4 都市行政評価ネットワーク会議年次報告書vol.1、 2006年、p19
- 5 同上4
- 6 総合研究開発機構自主研究『ベンチマーキング手法の地方自治体への導入』2003年、総合研究開発機構委託研究『ベンチマーキングの導入によるアーバン・マネジメントの改善』2003年、総合研究開発機構自主研究『都市自治体総合行政評価システムの構築に関する研究』2004年、日本都市センター自主研究『都市自治体の行政経営改革とベンチマーキング』2004年、総合研究開発機構委託研究『NIRA型ベンチマーク・モデルを活用した政策評価システム及び行政改善への提案に関する研究』2004年、総合研究開発機構自主研究『都市行政評価フォーラム(仮称)の構築に関する研究』2005年、総合研究開発機構委託研究『都市行政評価フォーラム(仮称)の構築の所究『都市行政評価フォーラム(仮称)の構築のためのデータ整備に関する研究』2005年
- 7 ©Microsoft
- 8 エクセル統計2006 for windows ©社会情報サービス (現在は2008を使用)
- 9 偏差値が導入できた場合は、47.5以上52.5未満を「標準的」とし、52.5以上55.0未満を「+」、55.0以上60.0未満を「++」、60.0以上70.0未満を「+++」、45.0以上47.5未満を「-」、40.0以上45.0未満を「-ー」、30.0以上40.0未満を「-ー」、70.0以上と30.0未満を「突出値」、正規確率プロット確認時点で外れ値となった場合を「外れ値」と評価。標準的を中心に、数値の高い側、低い側、それぞれが5段階となるため、全体で11段階評価。ヒストグラムでしか表せない場合も、平均値と中央値で標準域を設定し、偏差値同様に段階に分けて評価。
- 10「第4回保険者による健診・保健指導の円滑な実施方 策に関する検討会資料5 特定健康診査等実施計画 における目標値及び後期高齢者支援金の加算・減算 について」p4 (http://www.mhlw.go.jp/shingi/2007/ 01/dl/s0117-8e.pdf)
- 11 厚生労働省老健局総務課「公的介護保険制度の現状と今後の課題 平成22年」2.介護保険制度の基本的な仕組み

12『都市行政評価ネットワーク会議年次報告』vol.1 (2006年) ~vol.4 (2009年) 等

13 12に同じ。

14 平成22年12月17日閣議決定「第3次男女共同参画計

画」第2部施策の基本的方向と具体的施策 第1分野 政策・方針決定過程への女性の参画の拡大

(2011.8.9受理)

# Municipal Evaluation Consortium: Benchmarking for the Improvement of Local Government Service

Hiroyasu Kumakura

Center for Gunma Studies, Gunma Prefectural Women's University npo-kuma@xp.wind.jp

#### Abstract

The Municipal Evaluation Consortium was formed seven years ago to break the impasse that municipalities in Japan had been facing in their evaluation practice. Since the Consortium was established its membership has grown from fifteen cities to eighty. Activities of the Consortium now include the calculation of deviation values for twenty public service areas and the provision of analysis results, called *karte*, for each participating municipality.

In the analytic approach adopted by the Consortium, the following four types of performance measures are used; basic measures that represent the circumstances involving public services, results measures that reflect the results of public services and activities, outcome measures that show the degree to which citizens respond to or utilize public services and activities, and cost measures. The effectiveness of employing such measures is recognized by many public officials and employees involved in the Consortium.

In this article, the families of measures developed by the Consortium are presented for the following public service areas; adult health examination service, long-term care insurance projects, library service, community center administration and operation, fire and emergency service, garbage collection and disposal, and gender equality programs.

The analysis results, *karte*, that have been developed and refined by trial and error by the Consortium are explained. The cases where *karte* are effectively utilized are also presented. Finally, suggested improvements in public services and activities using benchmarking and further problems are discussed.

## **Keywords**

municipal administrative evaluation, benchmarking between municipalities, Voluntary participation of municipalities, four types of measures, "Karte"

## 【実践・調査報告】

## 川西市における行政評価の取り組みとベンチマーキングの可能性

## 本荘 重弘

兵庫県川西市役所総合政策部長 hs0407@city.kawanishi.lg.jp

## 要約

行政評価なるものに取り組み始めて早19年が経過する。この間、様々な試行錯誤を重ねてきたが、その結果、当初に意図した総合計画を基軸とした行財政運営の仕組みの構築と、それを支える職員の成果指向型意識の醸成という点においては一定の前進を見たのではないかと自負している。

一方で、市民との関わりという点においては、評価結果を踏まえて、市民と課題を共有し、議論を深掘りするには至っていない。(このような反省もあり、7月には阪神間の都市で初めて公開事業レビューを実施したが)。実施計画の策定、予算編成などの作業プロセスの可視化も含めて、取り組むべき課題であると認識している。

その一つのツールが、自治体間ベンチマーキングではないだろうか。すなわち、行政評価の結果は、自らの自治体の、言わば成績表として自己評価されるが、その結果が、他の自治体と比較して優位にあるのか劣位にあるのか、また、その理由は何かというさらなる分析と公開を通じて、行政内部はもとより、市民の側も、望ましい税金の配分と業務執行の適正性を判断することができる。本稿では、本市の取り組みを辿りながら、ベンチマーキングの可能性について考察したい。

## キーワード

情報の公開性、自治体間競争によるインパクト、協働性

#### はじめに

本市における行政評価の取り組みは、1992(平成4)年、第3次の総合計画(計画期間:平成5年度~14年度)の策定作業にまでさかのぼる。総合計画に限らず、新規計画を策定する場合、当該計画分野に係る現状の把握・総括・評価を経て課題を抽出するとともにめざすべき姿を想定し、その実現に向けた施策体系と事務事業を整理する。第3次総合計画の策定に際しても、同様の作業に取

りかかろうとしたのであるが、総合計画を体系的 に管理・評価するという発想そのものが欠落して いた当時の状況の中で、計画を評価するすべを全 く持ち合わせておらず、結果、基本計画に記載さ れた文章を、事務事業単位の文節で断片化し、そ れらを「完了」「継続」「未実施」の区分で分類す るほかなかった。

また、総合計画に対する認知度について全職員 アンケートを実施したところ、「言葉だけは聞い たことがある」という回答が「大体知っている」

日本評価学会『日本評価研究』第11巻第2号、2011年、pp.45-54

46 本荘 重弘

をかなり上回っており、「全く知らない」も10%、さらに、「総合計画書を全く読んだことがない」という職員が40%も存在し、残念ながらその中には、部長、課長も多く含まれるなど、日常業務の中における総合計画の位置づけの低さを目の当たりにしたのである。

計画の持つ性格として、一つには、行財政運営の指針、二つには、国や県、他の自治体との相互調整を図るための指針、三つには、市民・事業者・行政など、まちづくりの様々な主体が共有すべき指針としての役割がある。

しかしながら、先述したような実態の中では、総合計画に期待される機能は働くべくもなく、計画策定の際に費やした市民参加や職員参加のコスト、コンサルタント委託費など多額なコストに対する便益は限りなくゼロに近い。

ましてや、多くの市民参加を得て作り上げた総合計画であるにも関わらず、職員の認知度が惨憺たる状況であるということは、裏切りとも言える行為ではないか。「一体何のための、誰のための総合計画なのか」。これが当時覚えた実感であり、行政評価に取り組む嚆矢であった。

#### 1. システムの変遷

## (1)(財)地方自治研究資料センターとの共同研究

前節での問題意識を抱える中、(財)地方自治研究資料センターにおいても、基本計画の事後評価に関する研究が進められていることを知り、渡りに船との思いでご相談申しあげた。

当時、同センターの主任研究員は、現高崎経済 大学名誉教授の斎藤達三氏であった。氏の指導の 下、①基本計画・実施計画・予算編成の相互関連 性の確保を軸としたシステム構想、②実施計画の システム化の中における定数管理の一体化、③事 業予算化率、事業執行過程、事業費執行率、事業 進捗率、施策・事業の達成成果などの指標化等を 開発方針とした研究を行ったところである。

しかしながら、①事業の成果指標を設定するに 当たって、指標の妥当性の判断を優先し、成果把 握のためのデータの有無についてはほとんど無視 したこと、②執行管理・評価する対象事業の絞り 込みが不十分であったこと、③システム検討時に おいては事業別予算が未導入であり、計画情報と 予算情報の共有が図れなかったことなどの理由か ら、結果として全庁的なシステムとして稼働する には至らなかった。

#### (2) 細事業(事務事業)評価の試行

このような課題を踏まえ、1997年度には、翌年度からスタートする後期基本計画の策定に合わせて、細事業ベースでの評価を試行するとともに、財政課においても、基本計画の体系に沿って予算事業を整理し、事業別予算の本格導入への準備作業が進められた。また、1998年度には、庁内LAN「川西ハイパーネット」が稼働するなど、評価がシステムとして機能する環境が徐々に整ってきた。

この段階における問題点として、①指標の設定を最優先した結果、一つの細事業に対し、アウトカム、アウトプットという性格の異なる指標が設定された、②通常アウトカム指標は、複数の細事業(手段)によって達成されるケースが多く、細事業レベルのみに着目した評価には限界があることなどが浮かび上がった。

## (3) 小項目(施策)・細事業(事務事業)の2段 階評価

このような細事業評価の課題を克服すべく、 1999年度からは、第3次総合計画の施策体系の内、 小項目(施策)及び細事業の2つのレベルでの評価を行った。

すなわち、まず、「政策アセスメント」という 括りで、総合計画の体系の小項目単位で「基本方 針」を樹立するとともに、アウトカム指標を設定 する。次に「事業アセスメント」の括りで「基本 方針」に基づき事業・細事業を精査するとともに、 アウトプット指標・インプット指標を設定すると いうリフォームである。

また、2000年度には、一般会計の決算を総合計画の体系ごとに整理するとともに、行政サービスの成果を事務事業の決算額と合わせて表示した「決算成果報告書川西こころ街レポート」を作成した。これは、地方自治法第233条第5項の規定に基づく、議会への提出が義務づけられた決算資料

であり、本市でも従来は「主要な施策の実績報告書」としてまとめていたものである。(成果報告書については次のURLを参照。http://61.122.230.36/gyozaisei/zaisei/kessan seika/21seika.html)

このように、総合計画の体系別にコストを整理するとともに、評価情報と決算情報を合わせて表示することで、①一つの施策が複数の部局によって達成がめざされており、その施策分野にどれだけの投資がなされているのか、②事業コスト(直接事業費+人件費+公債費)に対してどれだけのパフォーマンスがあがったのかなどが一目瞭然となった。最近では、議会の決算審査において「こころ街レポート」の情報を引用しての質問が多く出される傾向があり、結果として、それに対応する管理職は、事務事業の適正な執行もさることながら、執行によってもたらされる成果を見据え、事務事業と成果の因果をしっかりと説明することが求められるなど、従来の予算・執行重視から決算・成果志向へと意識の変容が追られている。

さて、この段階における評価の課題としてあがってきたのは、アウトカム指標の中には、行政活動からだけでは決して達成し得ないものが多数存在しているということであった。例えば、「ごみの一人一日あたりの排出量などはその典型的なものである。これに係る行政の事務事業として、「ごみの減量に向けた啓発事業」や「分別収集あるいは事業者の協力なくして、排出量を抑制することは不可能である。もとより、突き詰めていけば、アウトカム指標の持つ性格から、ほとんどすべてのものが、行政努力だけでは達成し得ないと言えなくもないが、市民等との協働がより強く求められる指標については、その取り扱いを意識的に分類する必要性を感じたのである。

#### (4) 行財政システム全般の見直し

折しも、2003年度からは第4次の総合計画がスタートすることもあり、これに合わせて評価システムのリニューアルを検討した。具体的には、総合計画の施策体系である分野・施策・細事業の3つのレベルでの評価を行うこととし、分野にはベンチマーク指標を、施策にはアウトカム指標を、そして、細事業にはアウトプット指標をそれぞれ

設定した。

この内、ベンチマーク指標は、先述した行政の 努力のみでは達成が極めて困難であると判断される指標群を整理し、これらを市民とのコミュニケーションツールとして位置づけるものである。時、恰も、総合計画審議会の、議論の中でも、協働とパートナーシップのまちづくりを進めていくためには、情報と目標の共有化が不可欠であり、ベンチマークのような指標の必要性が指摘された。

またこうした評価の仕組みのリニューアルとともに、総合計画・中期財政収支計画・行政SR作戦<sup>2</sup>を本市の三位一体と捉え、行財政システム全般にわたる見直しを進めた。

### 2. わが国自治体の総合計画

このように、本市においては、総合計画の策定 を契機として行政評価の取組みを開始したところ であるが、そもそも、わが国における総合計画は、 どのような時代背景の下に生まれ、どういう役割 が期待されていたのであろうか。

#### (1) 総合計画の変遷

#### ①黎明期

自治体における総合計画の契機は、1953(昭和28)年からの昭和の大合併に遡る。すなわち、町村合併促進法においては、合併町村に「新町村建設計画」の策定を義務づけた。町村合併促進法は時限法で、1956(昭和31)年9月31日に失効することとなるが、新市町村のさらなる育成等を図るため、新市町村建設促進法が1956(昭和31)年6月30日に公布された。

同法の施行通達(1956(昭和31)年7月16日)では、合併時の新市町村建設計画について、「合併前においてそれぞれの立場を持ったここの市町村が協議して定めたものであるから、時には新市町村としての一体的な立場から再検討することを適当とするものがあり、またその実施に関する計画が必ずしも市町村の財政力に相応して的確に樹立されていないものもあるように見受けられるものもあるので、真に新市町村建設計画を合理的に実施するために必要に応じて再検討して、これを

48 本荘 重弘

調整又は変更するものとされていること」として いる。

ここから見て取れるのは、合併を機に、多くの 市町村が公共事業の拡大化を図ろうとしたことで あり、結果、身の丈を越えた建設計画が多く策定 されたのであろう。時は平成に移っても、同様の 現象が生じているのは皮肉なことである。合併関 連の補助金が1960(昭和35)年に補助金が打ち切 られると、計画は改定されず放置されることとな る。

こうした中、1965(昭和40)年、当時の自治省 行政局振興課は、市町村計画策定方法研究会を発 足させ、1966(昭和41)年3月末に報告を受けた。 この研究報告に基づく指導が、市町村の基本構想 の法制化と相まって、自治体における総合計画策 定が推進された。

#### ②市町村計画策定方法研究会報告の概要

研究報告では、従来の新市町村建設計画の基本計画が10年であり、前・後期各5年の実施計画と合わせて2段階方式であったのに対し、基本構想10年、基本計画5年、実施計画は3~5年のローリング方式の3段階方式を提唱している。

基本構想では、地域課題の解決を図った上での ビジョン、解決を図るための手段を内容とし、基 本計画では、基本構想に掲げられた諸施策のうち、 市町村が直接に実現手段を有する施策及びそれら 施策を合理的に推進するための内部管理合理化方 策の大綱について定めることとし、土地利用など 基礎的条件の整備計画、施設計画、非施設計画、 行財政計画について、財政と事業費の調整を行っ た上で作成することとしている。さらに、実施計 画については、事業計画と財政計画に分類し、事 業計画は、基本計画に掲げた事業の内、実施を決 定するものを決定のつど計上する長期的な予算と しての役割を、財政計画は、全体的な歳入、歳出 の見通しを立てるもので、事業計画に計上した事 業が実施できるか否かをチェックする役割を担 う。計画期間は3年とし、毎年度予算編成の前に 改定しながら後年度分を逐次繰り入れていき、常 に原則として3年先までのものが計上されている ようにするとされている。

現在における、自治体総合計画は、当時の研究

会報告の内容に概ね沿った形式で策定されている ものと思われる。

ここで興味深いのは、当時から、市町村計画の問題点として、(1)計画の性格についての理解が十分であるか、(2)現実の経済の動きや当該市町村を巡る社会的・経済的諸条件を無視し、計画の目標に妥当性を欠いていないか、(3)市町村の任務についての認識が不十分でないか、計画策定にあたって広域的配慮が欠けていないか、(4)生きた計画として現実の行政に役立っていないのではないかなどといった点が指摘されていたことである。

これらを克服し、基礎的な態度として、科学性、総合性、具体性の3点を挙げ、特に留意すべき点として、(1) 現状分析を十分に行うこと、(2) 現実に即した適切な目標を定立すべきこと、(3) 市町村の果たすべき役割を的確に反映した計画とすること、(4) 適切な計画の手法を研究すること、(5) 広域行政を計画に反映すること、上位計画等との関係に留意すること、(6) 計画作成のための体制を整備すべきことを挙げている。

さらに、自治体総合計画制度的に内包する困難性について次のように言及するとともに、それらは乗り越えるべきものと励ましている。

- ①地方団体を巡る社会経済的条件が、地方団体の 意向と関わりなしに変動し、かつ、地方団体がそ れに働きかけうる余地が少ない。
- ②計画実現手段で、当該団体が主体性を持ちうる ものが少なく、当該団体が投資主体とされている ものでも、財政的に国に依存しており、本当の意 味での主体性を持ち得ない。
- ③当該地方団体の区域について、国や都道府県の計画が重畳的につくられているが、それらとは別に当該団体の計画がつくる意義が乏しく、つくるとしても調整がつけがたい。
- ④計画をつくろうにも、基礎資料、統計資料が不備であり、また、計画策定にあたりうる人材が不足している。
- ⑤毎年の予算編成すら政治的にゆがめられがちであるのだから、現実に即して長期的な総合計画をつくろうとしても、政治的な混乱以外の何物をも生み出さない。
- ⑥格別に長期的な計画をつくらなくても、予算を

通じて地方団体の行政を計画的に運営することが できる。

地方分権や地方自治体における人材面での成長 など、今日的にはややそぐわない指摘もあるもの の、現在もなお、多くの総合計画担当者を悩ます 課題であると言えるだろう。

#### 3. 自治体総合計画の現状と課題

それでは、自治体総合計画の現状と課題はどうであろうか。2002(平成14)年2月に、(財)日本都市センターが行った「自治体の計画行政に関するアンケート調査」結果から概観してみる。

#### (1) なぜ総合計画を策定するのか

まず、どのような目的で総合計画を策定するのかという問いに対し、「総合的な観点からの政策の体系化を図ること」という回答が最も多く98.9%、次に、「自治体経営の将来見通しを明らかにすること」が95.6%、「市民への行政活動の説明をすること」が93.6%となっている。

計画的で体系的なまちづくりを進める上において、総合計画がその指針として位置づけられるとともに、市民に対して行政活動を説明するためのツールとして活用されている様子がうかがえる。

# (2) 計画と予算の連動はどのように行われているのか

では、計画の実効性を確保するために、予算との連携はどのように図られているのであろうか。

最も多いのが、「財政部門に予算要求書を提出するに先立って、企画部門と事業部門との間で協議をする」の36.7%、次いで、「予算要求書を作成する際には、各事業部門内限りで総合計画との調整に留意することとしており、それ以上に関係部門間で調整することはない」が18.9%、「財務部門のみで調整を行う」が12.8%、「予算要求が出揃った段階で、企画部門と財務部門とで調整のための協議を行う」が10.3%となっている。

このことから、ほぼ半数の自治体において、予 算作成のプロセスに企画部門が介在し、総合計画 との調整が図られていることがわかる。

## (3) 総合計画と評価の連携はどのように進められているのか

自治体における評価システムの導入の契機は 様々であろうが、総合計画との連携の度合いはど うであろうか。

都市自治体では、「実施計画、事務事業レベルの評価システムを導入している」ところが最も多く、17.0%となっている。次いで、「基本計画、施策レベルでの評価システムを導入している」とするところが8.6%となっている。また、僅かではあるが、「基本構想、政策レベルでの評価システムを導入している」ところも1.3%ある。

一方で、「総合計画との連携はしていない」とする自治体も6.7%存在する。「評価システムを導入していない」とする自治体が約5割ある中で、導入している自治体の約半数が、何らかの形で総合計画との連携を図っているようである。

## (4) 現在の総合計画は何を課題としているのか

総合計画を巡っては、これまでも総花的・網羅的であり、実効性に欠けるなど、多くの批判があった。

今回のアンケートでも、「事務事業の優先順位が不明確」と回答した自治体が最も多く71.7%、「事務事業削減のための方針として機能していない」が69.2%、「総合計画の内容が総花的になっていること」が67.1%、「職員に計画スケジュールや進行管理の意識が希薄である」が67.1%、「職員に計画の重要性が認識されていない」が59.2%、「マネジメントの視点に欠けている」が57.2%と高率になっている。

いずれも従来から問題視されてきた点であり、総合計画の形式や自治体行政における位置づけが、法施行後約40年を経過した現在も大きく変容していないことが分かる。

## (5) これからの総合計画はどのような姿が望ま しいのか

それでは、今後の総合計画のあり方としてどのような姿が望ましいと考えられているのであろうか。

最も多かったのが、「現行の総合計画システム に、予算、評価等を連携させること」で70.6%、 50 本荘 重弘

次いで、「3層構造を基調とし、政策の目的―手段 体系をより明確にすること」が57.0%、「施策等 の優先度を明確にする等、行政改革に役立つよう にすること」が40.7%であった。

一方、現行の総合計画を抜本的に見直すという立場から、「3層構造を必ずしも採用せず、目的一手段体系を明確にしたもの」という意見が18.4%、「計画期間を短期化し、社会経済状況の変化に柔軟に対応できるようにしたもの」という意見も18.4%の割合で指摘されている。

前項であげられた課題を克服するための具体的な解決策として、いずれも早急に検討し、実行すべき事柄である。

# (6) 職員は総合計画をどのように認識しているか

最後に、総合計画について、職員はどのように 認識しているのであろうか。

結果として、当然のことであるが、総合計画の 担当者は日々常に高い認識があるという回答となっている。また、財政担当部局については、必要 性に応じた認識、活用をしているが、事業部局で は、あまり高い認識・活用を行っていないという 回答となっている。

これらはいずれも、総合計画担当者の回答であり、実際に、財政担当部局や事業部局の職員の回答を想定した場合、さらに、意識の乖離が見られるのではないだろうか。

これら総合計画に対する職員の認識度合いも、 前項、前々項で挙げた総合計画そのものが持つ課 題を克服し、望ましい姿へと脱皮を図る中で変化 する、否、変化させなければいけない点であろう。

## 4. 総合計画と行政評価の展望

#### (1) 望ましい総合計画の姿

①アンケート等から見えてきたもの

前章で整理したわが国の自治体総合計画の変遷 及び現状等を踏まえて、今後における望ましい自 治体総合計画の姿について考察したい。

まず、総合計画を策定する目的として、多くの 自治体が挙げているのが「政策の体系化」という

点であった。確かに、今日、行政需要とこれに対応するサービス供給は多種多様であり、政策分野ごとに細分化・専門化する傾向にある。一方で、一つの行政課題の解決に際して、単一の政策分野の単一の施策のみで対応しうることは極めてまれであり、部門間の調整が不可欠な状況となっている。こうした中にあって、正しく、各部門における計画を総合化する作業が不可避となるのであり、その役割を担うのが総合計画であると位置づけられているのであろう。

しかしながら一方で、現在の総合計画では、「事務事業の優先順位が不明確」である、あるいは、「内容が総花的になっている」、「職員に計画スケジュールや進行管理の意識が希薄である」という課題を抱えているとする回答の高さは、「政策の体系化」をめざしながらも、実態との乖離の大きさを示唆するものである。

こうした課題を踏まえて、今後の総合計画のあり方としては、「現行の総合計画システムに、予算、評価等を連携させること」や、「政策の目的一手段体系をより明確にすること」、「施策等の優先度を明確にする等、行政改革に役立つようにすること」との回答が多かった。

自治体総合計画が抱える課題は多く、また、短期的には克服しがたい性格のものではあろうが、果たすべき役割や期待もまた大きいことを調査結果は示しているのではないだろうか。いずれにせよ、法定後約40年が経過した現在、総合計画のスタイルも大きく見直す時期にきていることは確かである。

#### (2) 本市の取り組みから

本市においては、「一体何のための、誰のための総合計画なのか」という問題意識が、行政評価の取り組みの契機であった。その背景には、そもそも総合計画は、行財政運営の指針という、いわば団体自治を担保し、強化するための計画という役割に加えて、市民やNPO、事業者、行政など様々なまちづくりの主体が共有すべきビジョンという、住民自治を形成し、実質化していくための計画という側面を併せ持っている。にもかかわらず、総合計画に対する行政職員の認知度や活用度が極めて低いという状況は、やや大仰に言えば、

地域における自治の営みを放棄することに繋がる のではないかという強い危惧があり、総合計画を 実質的に担保する一つのツールとしての行政評価 に期待をかけたという当時の事情があった。

以来、数次にわたる行政評価システムの変更を 行うと同時に、予算編成手法や庁内の意思形成シ ステムの見直しを進めるなど、行財政システム全 般にわたる改革を進めてきた。その結果、当初に 意図した総合計画を基軸とした行財政運営のしく みの構築と、それを支える職員の成果志向型意識 の醸成という点においては一定の前進をみたとこ ろである。

一方で、総合計画審議会に計画のフォローアップ機能を付加し、総合計画を広く地域におけるまちづくりの主体と共有し、各主体間を繋ぐメディア(媒体)として活用する取り組みは始めたものの、地域全体へのインパクトの度合いは、現状では質・量ともにまだまだ極めて低いと言わざるを得ない。

## 5. 地域におけるガバナンスと総合計画・ 行政評価

#### (1) 戦略的な地域経営計画へ

本市の第4次総合計画「川西こころ街計画2012」では、「新しい自治・まちづくりのしくみ」を構築しながら、「補完性の原理」によるまちづくりを進めることとしている。すなわち、統治秩序形成機能は専ら公共部門が担い、そこで形成された統治秩序を駆使して公共部門が社会的利益を増進していくという図式ではなく、社会を構成する多様な主体がその機能を担い、相互に補完し合いながら合理的な統治秩序を形成していくという方向性、すなわち「ガバナンス」をキーワードとした地域経営改革である。

このような地域におけるガバナンスの実現に向けて、総合計画、そして行政評価はそのツールとして有効に機能するのか。筆者は、条件を付して「YES」とお答えしたい。すなわち、総合計画、行政評価が、ともに地域における公共経営を支えるアイテムとして進化を遂げることである。

まず、総合計画については、事務事業積み上げ

型から地域戦略型への抜本的な見直しが必要である。多くの場合、基本構想に基づく基本計画を策定するとともに、予算編成を行う前段で毎年度実施計画を策定している。つまり「総合計画」というとき、構想といういわばビジョンレベルの記述のみならず、具体的な事務事業を整理した総体を称しているのである。

このような事務事業積み上げ型の発想がベースにある総合計画が、一般に「総花的」、「網羅的」と揶揄される所以である。もちろん、その背景には、一つには、従来から、総合計画は、議会や利害関係者の利害調整の受け皿としての位置づけがなされてきたこと、二つには、事業計画が、行政各部局からボトムアップ的に提出され、計画全体の調整が不十分なまま処理されているということ、三つには、事務事業を単位とする予算編成手法にあっては、総合計画計上事務事業が予算査定の重要な基準となるため、各部局が競って、自らの部局が所掌する事務事業を優先的に計上しようとする行動様式がある。

本市においては、総合計画と財政収支計画、行財政改革を三位一体となって取り組んでおり、完全なアナーキー状態にはなってはいないものの、やはり、事務事業を基軸に据えた構成となっている点は否めない。このようなスキームでは、いかに市民の参加を得て作りあげた計画といえども、官僚統制の呪縛から逃れることができず、「行政の計画」として矮小化されてしまう。これを回避するためには、市民、議会、行政など様々なステークホルダーの共同作業によって、地域の課題とこれに対する戦略的な対応方針を樹立し、方針に基づいたプログラム設定を行うという「戦略ブレークダウン型の地域計画」に改革することが必要である。

一方、行政評価のスタイルも、いわゆる行政の 生産活動の効率性・経済性・有効性の評価という 現在のスタイルは残しながらも、先の「戦略ブレークダウン型地域計画」と連動した、ステークホルダー間の共通言語となりうるベンチマークを多様な参加を得て作成・設定し、進捗状況をフォローアップするという新しいモデルが求められる (最も、このレベルに至れば「行政評価」という 表現そのものがフィットしないということになる 52 本荘 重弘

が)。

自治体の首長選挙においても、マニフェストを 掲げる候補者がさらに増加することが予想される 中で、総合計画や行政評価も、このようなドラス ティックな改革なくして、政治との連動を図るこ とが困難になる時期が早晩訪れるだろう。

#### (2) 行政経営品質向上プログラムの導入

こうした状況の中、本市においては、平成20年度から行政経営品質向上プログラムを導入し、「卓越した経営」をめざした取り組みを加速したいと考えている。

行政経営品質向上プログラムとは、顧客、競争、職員、社会のすべてにおいて高い価値を創造するために、革新し続ける経営づくりを支援する一連のプログラムで、米国の競争力復活の原動力の一つとなった「マルコム・ボルドリッジ国家品質賞」のセルフアセスメントの考え方を基点にしている。

では、なぜ、今行政経営品質向上プログラムを 導入するのか。答えは明白である。すなわち、市 役所は何のために存在するのかという「あるべき 姿」や組織としての「望ましさ」を継続的に、し かも自律的に学習し、「価値前提」の経営を行う ことが不可避であるからである。

総合計画と行政評価を結びつける取り組みによって、本市職員の成果重視志向は格段に高まったと実感している。しかしながら、住民起点の行政運営が徹底されているのか。各セクションの部分最適に止まっており、全体最適の追求がなされているのか。職員の自主性と創造性が遺憾なく発揮されているのか。などといった観点から現行の仕組みを総括したときに、必ずしも高い評価が得られる状況ではない。

そこで、これまでの取り組みをベースにしつつ、 行政経営品質向上プログラムがめざす基本的な価値、態度、信念、行動基準をしっかりと設定する とともに、職員自身がその実効性を検証する仕組 みを整えたいと考えている。

## 6. ベンチマーキングの可能性

本市では、2005(平成17)年に、総合研究開発機構主催の「都市行政評価フォーラムの構築に関する研究」の研究会メンバーの一員として参加した。この NIRA型BMの一連の取り組みについて、筆者は大いに評価し期待している。その背景には、先述したように、行政評価は、最終的には地域における公共経営の改革に向けての有効なツールとなりうるものであると捉えており、今回のNIRAの取り組みが一つの起爆剤になるのではないかとの予感があるからである。

その1点目は、情報の公開性、しかも、全国に向けての発信のインパクトである。

本市でも経験済みであるが、評価を職員間で動かしている間は、緊張感を欠き、しかもお手盛りの評価にならざるを得なかった。しかしながら、「こころ街レポート」の公表以降は、明らかに議会、住民の視線を意識した取り組みに変化した。正に「公開は力」であると感じている。

2点目は、自治体間競争によるインパクトである。

およそどのような組織においても、イノベーション、そして生産性向上への工夫は、すべからく競争による切磋琢磨から生まれる。ところが、ほとんどの行政機関は競争にさらされていない。こうした競争のない世界において、NIRA型BMによって浮き彫りになる地域と行政サービスの実態は、住民のイメージ形成において擬似的な競争環境を提供する。

そしてそのことが首長や行政マンの行動様式を 変える、あるいはイノベーションの契機となりう ると考えるのである。

3点目は「協働性」である。

2003年の12月には、市民活動団体や地域コミュニティの方5名に集まっていただいて、指標に関するヒアリングが行われたが、最初は、出席者の間に何をどこから議論していいのかという戸惑いがあったという印象を受けたが、次第に活発な意見交換が行われ、指標そのものの妥当性や足らざる部分の指摘もなされていた。

私は、この光景そのものが、BMの可能性を示唆していると感じている。すなわち、市民と行政

の共通言語としての機能をBMに持たせようとするならば、指標の設定時から多様なステークホルダーの参画が必要であり、従って、その評価についても多元的なまなざしが必要であると考えるからである。

本市の行政評価の取り組みの中でも、市民とともに評価をするという取り組みは緒に就いたばかりであり、また、その方法も改善すべき点が多い。今回のNIRA型BM、あるいは、都市行政評価ネットワーク会議についても、おそらくは一挙に理想型に到達するという性格のものではなく、まずは自治体行政内部の施策改善ツールとしての機能を充実させた上で、あるいは、同時並行的に充実させながら、最終的には多様なステークホルダーが共有し、的確に、また、冷静にデータを分析する中で、自らの自治体の姿を検証するという姿に到達するのではないかと考える。

今後の発展に期待を寄せるとともに、可能な限りの協力をさせていただきたいと考えている。

#### おわりに

本稿では、川西市における行政評価の変遷をご紹介するとともに、わが国自治体の総合計画が抱える課題を整理した上で、両者の望ましい姿について考察した。

その背景には、地域におけるより良きガバナンスを実現していくためには、両者を統合させた制度設計と活用が不可欠であるという筆者の強い拘りがある。

とりわけ、住民に最も身近な政府である基礎的 自治体においては、市民意思と行政意思、そして 議会意思が共有できる目標設定とその達成に向け たコ・ラボレイトとコ・プロダクトが必須となるのであり、こうした言わば「協働まちづくり」の仕組みを構築する上において、総合計画と行政評価は最も有効なツールであると確信するからである。但し、最終章において整理したように、両者が現在のような形式のままでは機能しないことは明白であり、本市も未だ道半ばにある。今後とも、さらに検討を重ね、あるべき姿を模索していきたい。

#### 注記

- 1 地方自治法第233条第5項「普通地方公共団体の長は、 決算を議会の認定に付するに当っては、当該決算に 係る会計年度における主要な施策の成果を説明する 書類その他政令で定める書類を合わせて提出しなけ ればならない
- 2 本市の行政計画の愛称。行政課題や日常業務を迅速 (Speedy) に処理することによって、市民の笑顔が あふれるような (Smile)、満足度の高い (Satisfaction) 行政を確立するため、現行の行財政 システム全般を見直す (Research) もの。

#### 参考文献

(財)日本都市センター (2003)『自治体と計画行政』 本荘重弘 (2000)「兵庫県川西市の行政評価の現状」、 『NIRA政策研究』、13(2)

本荘重弘 (2005)「行政評価システムを活用した行政 改善と今後の課題等」、『NIRA政策研究』、18(7) 日本経営品質賞委員会 (2008)『アセスメント基準書』 (財)行政管理研究センター (2008)『評価クォータリー』、5

(2011.8.12受理)

54 本荘 重弘

# The Administrative Evaluation of Kawanishi City and the Possibility of Benchmarking

Shigehiro Honjo

General Planning Department, Kawanishi City hs0407@city.kawanishi.lg.jp

#### **Abstract**

Nineteen years have passed since the Kawanishi City introduced and managed the administrative evaluation in 1991. After repeated trial and errors, the creation of an administrative and fiscal management system based on the general plan for Kawanishi City has been achieved. Progress has been made in building a strong objective oriented sense in the staff to support the newly created system.

On the other hand, there it is still difficult to share and explore issues with citizens through discussions based on the results of evaluation. In this connection a public review meeting on administrative activities in July 2010 was organized as the first trial among the cities in the Hanshin area. The formulation of a general plan, the preparation of a budget and the visualization of these activities are important issues to consider.

Benchmarking is one interesting tool for handling these issues. The results of administrative evaluations are assessed as report cards of the local government. The administrative staff and citizens judge the appropriateness of the budget allocation and its execution through the analysis and disclosure of the results by comparing them with other cities and identifying its advantages and disadvantage. The possibilities of benchmarking are analyzing through the trial of Kawanish City.

## **Keywords**

Kawanishi City, Administrative Evaluation, Benchmarking, Local Government, Budget Preparation, General Planning

## 【実践・調査報告】

## 福井市の行政評価と自治体ベンチマークシステム

## 中村 葉子

福井市総務部政策調整室 seityou@city.fukui.lg.jp

## 要約

近年、多くの地方自治体において行政評価が導入されている。福井市においても、平成11年度より行政 評価制度の導入に着手し、試行錯誤を重ねた過程から福井市独自の支援システムである自治体ベンチマー クシステム「比ベジョーズ」を構築した。その目的は、参加自治体間の相互提案によるよりよい成果指標 の追求と、相互比較による優れた自治体施策や業務改善への活用であった。

本論では、「比べジョーズ」の運用過程をたどりながら、研究事業などから見えてきた「比べジョーズ」の課題や成果を考察する。今回の分析により、自治体間比較の有効性を確認し、より実効性の高いシステムとして展開していきたい。

### キーワード

行政評価、自治体ベンチマーク、比ベジョーズ、特例市、アウトカム指標

#### 1. はじめに

行政評価(Performance Measurement)は、平成8年ごろから三重県や静岡県などの一部の地方自治体で先駆的に導入され、その後全国の自治体に急速に普及した。2009年の調査では、都道府県の91.5%が導入し、市及び23区では試行段階を含めて82.4%が何らかの取組をスタートさせている(三菱総研 2009)。

その背景には、地方の厳しい財政状況や地方分権の推進、自治体への不満等を契機とした住民意識の変化などにより、行政が成果や説明責任を強く求められるようになったという、自治体共通の課題がある。

自治体が導入する行政評価システムの手法は

様々であるが、実際の業務に携わる担当職員からは「期待したほどの効果がない」「職員の負担が増えるだけ」といった声が少なくない。それは、評価に膨大な事務負担がかかる割に、評価した内容が効果的にフィードバックされないことに現場の徒労感が募るからである。

平成11年度から事務事業評価の試行に着手した福井市も、同じように職員が「評価疲れ」をおこすような状態に陥ったが、総合計画の数値目標と連動した施策進行管理を導入することで、一定の成果を上げてきた。

本稿では、福井市における行政評価導入の経緯と、行政評価の支援ツールとして開発した自治体ベンチマークシステム「比ベジョーズ」の紹介、システム構築の中から派生した「比ベジョーズ研

56 中村 葉子

究会」の取組を報告する。そして、ベンチマーク システムを運営する中で直面してきた課題や成果 を整理する。

## 2. 福井市の行政評価制度の流れ

#### (1) 事務事業評価の着手

平成9年度の一般会計が赤字決算となった福井市は、財政赤字を克服するためのツールのひとつとして、平成11年度から実験的に事務事業評価の導入に着手した。導入にあたっては、三重県、東京都大田区及び埼玉県志木市などの行政評価手法を先進事例として研究し、参考とした。

その方法とは、行政改革を担当していた行財政 改革室が作成した評価シートを基に、個々の事業 ごとに絶対評価で点数化や優先順位付けをしよう とするものであった。しかし、採点基準があいま いであったため、各部局の点数差に大きな格差が 生じ、その結果、提出された評価シートは参考資 料としての信頼性を得られず、予算査定と連動さ せることができなかった。

他市の行政評価においても先導役を果たしているのは、ほとんどが企画部門や総務部門である。福井市の事務事業評価も、個別事業の効率化の検討は可能だが、事業の廃止といった大胆な予算の組換えは困難であるとの結論に至り、当初期待したような財政再建のツールとはならなかった。

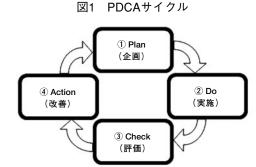
#### (2) 施策進行管理と行政評価支援システムの導入

行政評価は、経営改革のための業績目標と成果を測定する尺度にすぎず、業績や成果を向上させるための手段として活用されなければならない(大住 2005)。より実行性のある行政評価制度とするため、平成14年度から、事務事業評価システムは政策調整室に移管された。

事務事業には①企画(Plan)②実施(Do)③評価(Check)④改善(Action)の流れがあり、また①企画に戻るという循環の過程がある(PDCAサイクル)(図1)。政策調整室では、これまでの行政評価は③評価の段階で終わり、④改善にいたるプロセスが見えにくいのではと考え、これをシステム化することに取り組んだ。

また、これまでの事務事業評価がうまく機能しなかった反省点のひとつに、それぞれの事務事業を個別に捉えて業務改善を検討していたことが挙げられた。市として上位の目標を達成していくには、各課の業務を体系化し、全体を見渡すマクロの視点を持って他の事務事業と比較することが必要だと考えた。

そこで、政策調整室では「評価から改善へ歩みを進めること」を最大目標として、施策進行管理システムを導入した。市の基本目標を整理するために施策を設置し、事務事業はそれを達成するための手段とした。評価制度は、事務事業の効率化を図るための業務改善ツールに位置づけを変え、予算編成の参考資料とするのではなく各課の自己点検ツールとし、評価シートの内容も大幅に簡略化した。



(出所) 筆者作成

平成14年度に行った施策進行管理に提案された 改善方策130件のうち、翌15年度で対応したのは6 割の77件に上った。

また、同年度からスタートした「21世紀を拓く 創造プラン(第五次福井市総合計画)」において、 施策と事務事業に整合性を持たせ、各施策推進の ために本格的に行政評価の取組が始まった。さら に、総合計画の施策にはじめて数値目標を設定し、 その数値の進行管理を通して、総合計画の達成状 況を検証することとした。

自治体が住民に情報公開や説明責任を遂行しようとするときに、まず必要なことは、事務事業の成果を明らかにすることであり、総合計画に定められたまちづくりのための施策の進捗度を認識することである(石原 2005)。行政評価は、事務事

業や施策の、優先順位や劣後順位を明確にする上で、必須のツールといえよう(石原 2005)。

こうした行政評価や総合計画における取組を進めていくうち、行政評価に関するデータを集中して管理・蓄積していない、他市との比較データやアンケートによるデータ収集が不足しているなど、いくつかの問題が出てきた。

そして、これらの問題を解決するため、行政評価を支援する3つのシステムを開発することになった。アンケートの調査を支援するシステム「聞きジョーズ」、総合計画に関する数値指標を中心に集中的に管理するデータベース「調ベジョーズ」、そして他市とのデータ比較を容易にする「比ベジョーズ」である。これらの行政評価支援ツールを総称して「ジョーズ3兄弟」と呼んでいる。

# 3. 自治体ベンチマークシステム「比ベジョーズ」の展開

#### (1)「比ベジョーズ」の立ち上げ

ベンチマーキング(benchmarking)は、民間企業では他社の優れたところを学び、それを基準にして自らの業務や経営を改善するという経営手法として用いられている。

また、自治体の行政評価においては、他の自治体と比較することで現状把握のツールとなっている。しかし、これまでは数値指標がどのように推移したかという「自治体内での経年の比較」が行われることは少なかった。なぜなら、それぞれの自治体が独自に他自治体のデータを収集するという作業は、大変な手間がかかる上に、照会される自治体にも負担がかかるからだ。しかも、複数の自治体のデータ収集の取組を継続的に行い、蓄積していくことは困難だった。

しかし、福井市では、数値目標を盛り込んだ総合計画の達成状況を検証する視点のひとつとして、他自治体との比較が必要であるとの認識に至っていた。

そこで、平成14年、福井市は他市と数値指標の 比較を容易にするための自治体ベンチマークシス テム「比ベジョーズ」の構築を、人口20万人以上 で、福井市と類似規模である特例市に対して呼び かけた。

その目的は、

- ① 指標を相互比較することで、自治体の相対的 位置を知り、より良い政策立案や業務改善に 活用する
- ② 施策の成果を含め市民に分かりやすいアウトカム (成果) 指標 を調査研究する というものだった。

この呼びかけに対して26市が参加に同意し、平成15年4月から「比ベジョーズ」の運用を開始した。その後も、特例市の認定を受けた自治体にはその都度参加勧誘を行い、また中核市に移行した自治体についてもそのまま参加できる体制にしたため、現在は特例市を中心とした34市の協力を得て、各自治体のデータを収集している(平成23年4月現在)。

#### (2) システムの概要

「比べジョーズ」は、福井市がサーバを持ち、データベースの管理を行っている。自治体ごとに発行しているIDとパスワードを入力すれば、参加自治体の職員は誰でも「比べジョーズ」にアクセスすることができるので、担当する企画部門の職員だけではなく、全職員が自由に閲覧・入力できるようになっている。

参加にあたってのシステムの経費負担は不要だが、定期的にデータ入力の義務を負うという合意 書を交わしており、年2回程度のデータ入力を参加自治体へ依頼している。そして、年度末には年 次報告書を作成して、参加自治体に配布している。

開始当初は事務局と各自治体が電話や電子メールを通して連絡をしていたが、「参加自治体間でスムーズな情報交換ができる環境を整備してほしい」との要望があり、平成17年にはシステム内に「比べジョーズ」担当者専用の電子会議室(以下、会議室という。)を設置した(図2)。この会議室で投稿した内容は指定したメールアドレスで受信することが可能で、システムに関する連絡や情報共有のほか、市政全般に関する照会や回答をはじめとする情報のやり取りなど、「比べジョーズ」内の指標と直接関わらない照会にも活用されている。

各参加自治体が入力した指標は分野ごとに表形

58 中村 葉子

### 図2「会議室|画面



(出所) 比ベジョーズシステムから

### 図3「比ベジョーズ|閲覧画面(1)

4 . + .	3	(F)	4	1 8	2	READ	8	3	4
RS BD	中止	更新	#-L		t# 15	READ .	35-17		y-% ≥840 U
自治体ベンチマークシステ		46.0							
比べジョー	っし								rozu X
		IN BA	FARIS IS	Ullestri	データ人	n land	C CHOICE	TOP ME	NU IDDOZE
データ制質									
平成10 💌		< 6 to 00 6	きクリックで有	使取得づち	つへ、自治化	おおうりゅう	アでその自治	(体の)49が表	BEST.
SHE SHAPE SS		環境							100
ONLY >>	- 61						環境基		-
OEE					環境[大	環境[大	連(水 質)生活		
<b>●観光</b>			市民1人		例1二酸	知二酸	環境項	低公害	公害苦
<b>●数</b> 用	3	指標	当たりの	ゴミの資源化率	化硫黄 (SO2)	化望克 (NO2)	目 生物 化学的	車導入	情受付 件数
PER			みの量		年間平 均値	年刊平 均值	放弃要 宋量	4	1+8X
●佐疫	-				A/18	A/18.	(BOD)(7)		
●介護 ●延祉							75%値		
ens		単位	g	. 56	ppm	ppm	mg/l	36	件
969	2	養津日	HIB331	H18331	H18331	H18331	HI 8.3.31	Ht 8.3.31	HI 8.3.31
●公服. H	1	A市	-	-	-	-	-	-	-
<b>0</b> 建铁	2	8 市	-	-	-	-	-	-	-
●上下水道 ●市民争商活動	3	C市	-	-	-	-	-	-	-
05R13	4	DH	-	-	-	-	-	-	-
<b>中心手事</b> 器	5	F 75	-	-	-	-	-	-	-
	6	FR		-	-	-	-	-	-
CSV977/0-F									

(出所) 比ベジョーズシステムから

#### 図4 「比べジョーズ | 閲覧画面(2)

■環境	
●指標名	ゴミの資源化率
●単位	%
●基準日、備考	H18.3.31
●算出基準、根拠、出典等	(所管課題) 資源ごみのうち再利用業者に引き渡した割合(重さ)逆有償等含む 調査問題 毎年

	自治体	96	基準L備考	グラフ
	全体平均	17.4		
1	A市	17.8	- 1	
2	B 市	13.6	- 1	
3	C市	15.3	- 1	
4	D市	17.5	- 1	
5	Εħ	12.6	- 1	
6	F市	15.5	- 1	

(出所) 比ベジョーズシステムから

式で表示し、さらに見やすいように入力指標ごとに棒グラフを作成している(図3、4)。また、蓄積したデータはエクセル形式で年度ごとにダウンロードすることにより、加工・分析が容易にできるようになっている。各事業の所管課が、事務事業に関するデータを取り出し、企画や業務改善のツールとして用いる活用法も想定している。

#### (3) 指標の選定

システム構築にあたっては、収集する指標の選定基準が大きな問題点となった。「比べジョーズ」は、自治体に関する様々なデータを網羅したデータベースではなく、各市が統一的な基準によってデータを提供し、かつ、各分野の状況を把握できるシステムを目指したものである。そこで、指標の選定にあたっては、「比べジョーズ」参加全自治体に対して2回の指標意向調査を実施し、さらに細かい調整を電子メールでやり取りしながら決定した。

数値指標の比較は同一の基準で行うことが基本だが、各市の扱っているデータには算出の根拠や 基準が多少異なる場合があったため、お互いが希望をよくすり合わせ、次のように指標選定のルールを定めた。

- ①大多数の自治体が調査可能で、採用を希望したもの
- ②各都市の状況を見るうえで、基本となるもの
- ③市民に分かりやすいもの(市民にとって身近なもの、全体の数値を単位あたりの数値に加工し、比較しやすくしたもの)
- ④長期間継続してデータの取れるもの
- ⑤各政策分野の状況を把握できるもの

運用当初は17分野135指標からスタートした。 各自治体が責任を持って指標を管理できるかを考慮して採用したため、結果として、福井市が総合計画で設定した数値指標であっても、条件に合わない指標は採用を見送った。

#### (4) 指標の見直し

実際に運用を開始してみると、収集している指標に対して「調査に労力のかかる指標は避けるべき」、「『量』より『率』を測る指標を」という意見が参加自治体から出てきたため、数値指標の調

査研究の必要が出てきた。

統一した基準で各分野を把握できるように選定した「比べジョーズ」の指標だが、さらなる活用促進をはかるため、データの検証可能性・客観性・有用性といった特性に留意して整理することにした。

協議の結果、ベンチマークに結びつくアウトカム (成果) 指標を中心に据えるが、指標が設定できない分野には、アウトプット (活動) 指標を工夫して置き、全体的なバランスを取ると定めた。また、バックデータとなるような最低限の基礎指標も入れることにした。

そこで、先の選定ルールに次の条件を加え、具体的な指標の選定に取り組んだ。

- ①人口関係の指標や財政関係の指標など、後々 指標を変換するときに活用できる指標は残す
- ②全自治体ができるだけ毎年調査可能な指標とする
- ③基本的に人口割のような割合の形に整理する
- ④法改正や制度改正によって新しく変わった指標については新しいものに変更する

平成19年度、後述する研究会において、数値指標の大幅な見直しに取り組んだ。各参加自治体から分野ごとに提案された指標案を一つひとつ丁寧に検討した結果、現行の135指標から、より自治体間比較に相応しいと考えられる165指標にまとめた。その後、全参加自治体への書面確認を経て、19年度分の入力から、17分野155指標で数値を蓄積していくことになった(図5)。

この一連の見直し作業を通して、数値指標を一 定の基準の下に整理し直したが、分野間のバラン スをとる難しさや、算出基礎の統一などの課題は 依然として残っている。

## 4. 「比べジョーズ」を利用した研究事例

#### (1) 研究会と分科会

こうして「比ベジョーズ」を運用していく中で、自治体間比較の試みを共同で研究するという新しい動きがでてきた。平成16年、システムの利用促進や指標の見直しなどについて協議・検討を行うため、福井市において研究会を開催した。その中で、収集している指標をどう活用したらいいのか、ベンチマーキングのやり方がわからないなどの意見が出た。そこで、希望する自治体間で、指標を比較して明らかになった課題について、分析や改善方策の立案を行う分科会を設置することが提案された。福井市でもシステム活用の最終到達点は事務事業の改善方策の立案とその実行としており、より詳細な指標の分析には参加自治体間の連携は必要不可欠であった。

そして、平成17年以降は研究会の下に分科会を設け、分科会における調査研究成果を年度末に開催する研究会にて報告することとした。以下では、これまでに開催した分科会及び研究会の中から、平成17年度及び20年度の取組を報告する。

## (2) ベストプラクティス都市の選定【平成17年度】 平成17年、第1回の分科会開催を呼びかけたと ころ、福井市を含め、7市の参加があった。自治 体間比較をどのように行うか、まったくの手探り

107 E	「 LL	ゔ゠	☆ 45 15 1年	臣生
図5	比ベジョ	ースー	登録指標-	一 見.

分野	指標数	分野	指標数	分野	指標数
1. 人口	14 (13)	7. 医療	7 (8)	13.建設	6 (5)
2. 財政	13 (10)	8. 介護	9 (11)	14.上、下水道	8 (5)
3. 産業	9 (9)	9. 福祉	6 (2)	15. 市民参画、活動	10(7)
4. 観光	2 (2)	10. 防災	14 (10)	16. 市民生活	2 (3)
5. 教育	27 (26)	11. 環境	7 (7)	17. 内部事務	5 (2)
6. 子育て・保育	11 (9)	12. 公園、緑	5 (6)	合 計	155 (135)

60 中村 葉子

状態であったため、民間企業のベンチマーキング を応用して、参加する自治体内でもっとも優れた 業績の市(ベストプラクティス都市)を選び、そ の優秀市の事業プロセスを分析し、優れたパフォ ーマンスを生み出す要因を特定することとした。

大住莊四郎 (2005) によれば、公共部門で用いるベンチマーキングの理想型のパターンは次のとおりである。

- ① 自らの組織の課題を発見し、課題を解決する ために外部のカウンターパート (外部の類似 組織など)とパフォーマンスを比較する。
- ② 「ベストプラクティス」を見出すことができれば民間企業でいうベンチマーキング・プロセスを活用し、自組織のパフォーマンス向上を図る。

テーマは、各自治体のサービス規格が比較しや すい「保育サービス事業」に決定した。

「比べジョーズ」に採用されている保育関連の 指標を抽出したが、わずか9指標しかなく、ベン チマーキングによる分析を行うためにはこの指標 だけでは不足していることが分かった。そこで、 所管課等にも確認しながら、新たに関連する基礎 指標の抽出と、併せて各市の保育サービスの現状 及び課題をまとめた。

そして、参加自治体間で基礎指標の簡易比較や、 保育サービス事業の状況及び政策体系における位 置づけなどについて情報交換を行い、集まった指標を上手に加工することで詳細な自治体比較ができないだろうか、という考えにまとまり、分析することとした。

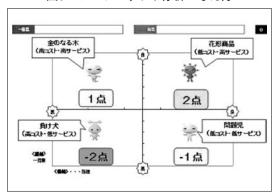
はじめに、「比べジョーズ」内の9指標と新たに 集めた43指標を、「基礎指標」、「インプット(コ スト)指標<sup>3</sup>」、「アウトプット(活動)指標」、 「アウトカム(成果)指標」に整理、分類した。

この作業の中で、1つの指標で自治体比較をすることも有用だが、いくつかの指標を複合的に比較することで、自治体ごとの違いがより明らかになることも分かってきた。そんな時、民間企業が用いるポートフォリオ分析を応用してサービスの水準比較分析をしてはどうかという提案を参加自治体から受け、この分析をしてみることにした。

ポートフォリオ分析の手法はいくつかあるが、 広く利用されているものの一つを採用した。"市 場成長率"と"マーケットシェア"の2つから、自社商品を花形商品(star)、金のなる木(cash cow)、問題児(problem)、負け犬(dogs)の4つに分類して有望な分野を探すという方法である(図6)。これを応用して、それぞれの自治体の強み・弱みを把握するとともに、低コスト・高サービスの部分に分類されることが多い自治体を探そうというものである。

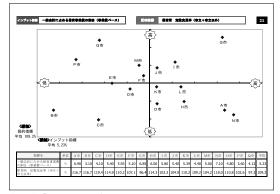
先に整理した指標のうち、コスト面からの指標とサービス面からの指標をクロスさせて、散布図を複数作成した(図7)。そして、花形商品エリア(低コスト・高サービス)に位置する自治体に+2点、金のなる木エリア(高コスト・高サービス)に+1点、問題児エリア(低コスト・低サービス)に-1点、負け犬エリア(高コスト・低サービス)に-2点を入力して、総合得点を算出した。

#### 図6 ポートフォリオ分析の考え方



(出所)「比ベジョーズ研究会」作成

#### 図7 散布図(見本)



(出所)「比ベジョーズ研究会」作成

この点数化の作業の中で指標を評価する場合、市民からの視点と行政からの視点のどちらに重心を置くかで、「良い・悪い」の評価が変わってしまうことがあった。例としては、「休日保育を実施している保育所の割合」が高い場合、市民にとっては保護者の多様なニーズに応えているという点で「良い」であり、行政にとっては保育士の配置負担や人件費がかかる点で「悪い」という評価になってしまう。このときは、市民からの視点で捉えるよう散布図の配置を修正した。

総合得点はかなりのばらつきがあり、点数が高く出た自治体、低く出た自治体ともに驚きの結果となったが、これを踏まえて、各自治体が保育サービス事業の特徴などの自己分析を行い、互いに報告をした。

こうして他の自治体の優れた取組を学び、それを各自治体に最も適した形で取り入れることで、大きな改善に結びつけることができるという認識を持った。

平成18年度は、同様の手順で「清掃事業」におけるベストプラクティス都市の選定に挑戦したが、平成17年度とは逆に、行政にとってのベストプラクティスを探るという視点に統一してまとめた。

この2年間の取組を通して、自治体における効果的なベンチマーキングの手法を検討したが、自治体が直接コントロールできない要素(地域性、規模など)もあり、一朝一夕に確立できる作業ではなかった。

しかし、これからの自治体経営を考えたとき、自分の自治体だけを考えていては新しい施策や事業は生まれてこない。他の自治体の状況を知らずに職場内部で業務改善について検討するよりも、他市と比較することでより早く改善のヒントを得られ、効率的な運営に繋がるのではないかと分かったことが、分科会の大きな収穫であった。

## (3) 仮想都市「比ベジョーズ市」の総合計画 【平成20年度】

「比べジョーズ」をより実用的に運用しやすいシステムとするためには、参加自治体がベンチマーキングに利用しやすい指標を取り入れる必要があるという考えを引き継ぎ、20年度の分科会では

「総合計画における数値指標の比較と設定」をテーマとした。

これは、「比ベジョーズ」が特例市を中心に、 ほぼ同規模の自治体が集まっているという強みを 活かし、参加自治体の総合計画を最大公約数的に 整理・比較することで、類似した自治体に共通す る課題や方向性を探るためである。

まず、参加自治体の総合計画の政策や施策の体系をつかむため、各自治体の総合計画を見比べた。体系を類型化するため、政策を「健康福祉」「教育文化」「環境」「産業」「都市基盤」「安心安全」「市民参画・行政経営」の7分野に整理し、さらにその下の施策を32分野に分類した。たとえば、「健康福祉」という政策は、「健康」「高齢者」「障害者」「子ども」「社会保障」の5施策で構成している。

さらに、特例市全体が目指そうとしている標準的な姿として仮想都市「比べジョーズ市」を設定し、参加自治体の施策レベルの指標と目標水準を積み上げることで、「比べジョーズ市」総合計画の策定を試みた。

選定する指標については、32分野に分類した施策の中に、参加自治体の総合計画で管理されている指標及び「比ベジョーズ」で運用している指標を当て込んでいった。「比ベジョーズ市」総合計画に採用した55指標の選定をするにあたっては、全自治体で把握可能かつ長期継続してデータが取れること、自治体の努力で解決できることなどに留意した。

また、目標水準の設定については、自治体の平均をとる方法、ベストの値をとる方法、自治体の上位25%をとる方法の3案が示された。しかし、平均をとる方法は半分の自治体が目標をクリアしてしまっており、また、ベストの値をとる方法は突出した値を目標とした場合達成が難しいことから、イギリスのベスト・バリュー(Best Value)の事例を参考に、上位25%の値を取る方法を目標水準と定めることとした。特例市の平均モデル上位25%を目標値にすることで、自治体がワンランク上を目指すために取り組みやすく、そして一般市民などから見ても分かりやすい目標水準が設定できた。

策定した「比ベジョーズ市」の総合計画は、市

62 中村 葉子

政運営の方向性と、数値指標を統一していくための一定のルールを示すことができた。算出した目標水準は、SWOT分析。などを独自に組み合わせることで、分野ごとの弱点や施策別の進捗度がわかる有効な尺度になり、各自治体の目指す方向、優先度・重要度を柔軟に判断する上で十分活用できるものとなっている。

### 5. これまでの成果と課題

福井市の行政評価支援システムのひとつとして 開発した「比ベジョーズ」の導入から7年あまり が経ち、本システムは17分野155指標のデータを 蓄積するデータベースと、研究会及び分科会の開 催という2本柱を中心とした運用が定着している。

本システムの最大のメリットは、類似規模の自治体が集まる中で、自分の自治体の立ち位置を容易に確認できることである。自治体として目標は立てたが、果たして同じグループの中でその目標がどの程度の水準にあるのか、他自治体との比較に基づくことで達成度を高めていくことができる。

導入当初は、参加自治体はシステムコストがかからず、自分の自治体の数値を入力すれば、加入する30以上の自治体の数値が閲覧できることから、それぞれの自治体でマネジメントや評価等に利用してもらえれば、というスタンスであった。

しかしながら、当時は行政評価などに取り組む 自治体も少なかったため、共同で利用方法を研究 する活動が自然に生まれてきた。福井市も行政評 価を導入する際、他の自治体の手法を安易に取り 入れても適切に機能しないという教訓を得てい る。分科会活動においては、他の自治体とベスト プラクティスの抽出や活用方法を共同で検討する ことにより、各自治体に適した行政評価の手法を 見出すことや、行政評価及び政策立案に携わる職 員の意識改革につながる役割を担っている。いま や職員のレベルアップに欠かせない有用な取組と なっているといえよう。

また、「比ベジョーズ」は自治体職員間のネットワークづくりに大きく寄与している。「比ベジョーズ」担当者が利用できる会議室では、通常、各自治体に照会をかけるような内容のものや、業

務上の参考として聞いてみたいものなどを、会議室に投稿することによって各自治体からの回答を、簡単かつ迅速に入手することが可能である。各自治体共通の課題や問題に対して、迅速に情報共有できる手段として有効に活用されている。

一方、「比べジョーズ」の運営をしている中で、 いくつかの課題も浮かんできている。

まず一つめは、研究会や分科会にメンバーの固定化が起こっており、ベンチマーキングのノウハウの蓄積に偏りがあることである。システムをいかに有効に活用していくかは、それぞれの自治体の意気込みや各自治体のマネジメントの考え方によって大きく左右される。分科会に参加した自治体は、他市との違いを知ることで再度自分の、そのノウハウを知ることができる。しかし、参加自治体は全国に散らばっており、研究会や分科会の参加が難しい自治体も出てくる。この課題を解決する方策として、前述の会議室への投稿や、研究会の報告書の送付により、情報の共有化を図っている。

二つめは、導入当初からの目的である「よりよいアウトカム(成果)指標の追求・選定」作業はまだ不十分であり、研究の余地があることだ。平成19年度に指標の全体的な見直しをしたが、さらに、毎年若干の修正を加えながら、よりよい指標の選定を目指している。ベンチマークに適切な指標を設定するには、今後も自治体間の議論を重ねていく必要がある。

また、指標の正確性についても、自治体によって算出基礎や根拠に微妙な違いがあるとの指摘があり、統一された測定マニュアルを作ることが必要である。そのためには、厳密な定義づけが可能なアウトカム(成果)指標を選定していくことが重要であるが、観光や福祉などの分野においては、成果を指標化しにくいという問題がある。

最後に、「比べジョーズ」の有効活用が進まない原因として、「比べジョーズ」のデータを非公開としていることがあげられる。「比べジョーズ」の構築当初、「都市ランキングとして興味本位に利用されるのではないか」と危惧する自治体への配慮や、フリーライダーの排除を理由に非公開とした。しかし、非公開であるゆえに市民に対する

情報提供のソースとして使いにくく、参加自治体の積極的な活用に繋がらないため、業務改善ツールとしての機能を十分に発揮できていない。ベンチマークの有用性・利点を広く啓発し次のステップに進むには少しでも外部利用できた方がいいという意見も出ており、「他自治体との比較情報」をどのように取り扱うべきか、引き続き検討する必要がある。

#### 6. おわりに

以上、自治体ベンチマークシステム「比ベジョーズ」の基本的なフレーム、研究会や分科会の概要、その成果と課題等を整理してきた。

「比べジョーズ」は、今後どのように活用していけばいいのだろうか。「比べジョーズ」の当初の目的は、効率性やサービス水準を他の自治体と比較してより優れた業務改善に繋げること、よりよいアウトカム(成果)指標を追及することであった。各自治体において自治体改革の方向性は必ずしも同じではないが、自治体には財政の窮迫や地方分権の推進など共通の課題がある。行政評価の手法で個々が抱える課題を明確かつ具体的に洗い出し、それぞれの求める視点で、自治体間比較の分析を応用すれば、「比べジョーズ」はより有効に機能するシステムとなるだろう。

「比べジョーズ」は、自治体のシステム改革や職員の意識改革を実現できる可能性を秘めている。今後は、「比べジョーズ」の参加自治体が主体的に分析に取り組み、互いに切磋琢磨し合えるような仕組みに発展するよう期待する。

そのためにも、福井市はベンチマーキングについての研究会や分科会の開催、会議室を通じた取組など、有効な活用策を発信し、より多くの自治体の参加を呼びかけていきたい。

#### 謝辞

この場をお借りして、「比べジョーズ」参加自 治体の皆様、そして関西学院大学の石原先生に謝 意を表したい。

#### 注記

- 1 行政活動等により地域社会に生まれた変化を示す指標。例)手話ボランティア登録者数、ボランティア 活動実績。
- 2 どれだけ行政活動が行われたかという実績を示す指標。例)手話ボランティア養成講座受講者数、講座 開催数。
- 3 行政活動に対して、どれだけ税金や労力を投入した かを示す指標。
- 4 2000年4月、イギリスにおいて、地方自治体の行政 サービス近代化を図るために導入された。自治体は、 作成したベスト・バリュー業績計画に基づきなが ら、5年以内に現在における全国レベルの上位25% 以内のサービス水準達成を求められる。コスト効率 においては、同格の自治体の上位25%と同等に成果 を達成すること及び歳出ベースで年率2%のコスト 改善を実現することとされている(努力目標)。
- 5 Strength (強み)、Weakness (弱み)、Opportunity (機会)、Threat (脅威) の頭文字を取ったもので、自分たちの組織の強みと弱みを分析するという意味。SとWを分析し、そのうえでOとTを考える。

#### 参考文献

石原俊彦編(2004)『新行政経営マニュアル イギリス のNPMに学ぶ』、清文社

石原俊彦編(2005)『自治体行政評価ケーススタディ』、 東洋経済新報社

上山信一 (2002) 『日本の行政評価』、第一法規出版 大住莊四郎 (2005) 『NPMによる経営革新』、学陽書房 三菱総研総合研究所 (2009) 『地方自治体における行 政評価等への取り組みに関する実態調査: 2009年』、 http://www.mri.co.jp/

(2011.8.13受理)

64 中村 葉子

## Fukui City's Performance Measurement and Municipal Benchmark System

Yoko Nakamura

Fukui City, General Affairs Department, Policy Coordination Division seityou@city.fukui.lg.jp

#### **Abstract**

Fukui City introduced a performance measurement system in 1999. To support this system, Fukui City developed its own municipal benchmark system called *Kurabe Jozu*, as a result of trial and error. The *Kurabe Jozu* is a more useful outcome index because it accepts proposals from participating municipalities and compares existing indices from other municipalities for better municipal measures and operational improvement.

The problems and outcomes of *Kurabe Jpzu* are discussed in the body of this report through research projects following its operation process. The effectiveness of *Kurabe Jozu* is confirmed and this analysis suggests further improvements.

## **Keywords**

Performance Measurement, Municipal Benchmark, *Kurabe Jozu*, outcome index

## 【研究論文】

## 行政部門評価におけるベンチマーキングの現実機能 一評価主体の協働と評価性向試論一

西出 順郎 岩手県立大学 nishide@iwate-pu.ac.jp

## 要約

本稿の目的は、行政ベンチマーキングにおける行政部門内外の評価主体が依拠する理論的枠組を解きほぐしながら行政ベンチマーキングの現実機能について検討することにある。

検討の結果、複数の理論的枠組が導出する評価主体の評価性向から行政部門評価では促進的もしくは調整的役割を果たす協働が有意的行為であることを明らかにし、同時に行政部門の長たる首長のオーナーシップが協働の内実化を妨げる評価実施上の脅威の源であることを確認した。その上で行政ベンチマーキングが当該脅威を低減させ、自らの現実機能に資するための実務的示唆について論及した。

## キーワード

ベンチマーキング、評価、協働、評価主体、評価性向

#### 1 はじめに

近年、中核市や特例市など同規模の都市自治体間でコンソーシアムが立ち上がり、ベンチマーキングによる行政品質の向上取組が展開されている。地方自治体のこのような行政部門のベンチマーキング(以下、行政ベンチマーキング)は学術領域でも注目され、特に管理研究よりも評価研究において行政部門の評価枠組のひとつとして積極的に議論されている。

評価としてのベンチマーキングは複数の利害関係者に同時かつ能動的に評価作業へ参画させ、協働を要請する。それ故にベンチマーキングは行政部門評価が直面する問題を低減させ得る長所を備えている。しかし当該協働の実現は、一連の評価

作業に能動的に関与を求める「利害関係者」(以下、評価主体) および彼らが期待する「評価結果の性向」(以下、評価性向)、それら両者の多様性に鑑みても決して容易なことではない。

本稿は行政ベンチマーキングの現実機能に資することを目的に多様な評価主体による協働について理論的な検討を試みるものである。

構成は以下のとおりである。はじめに行政部門のベンチマーキングについて概観する。そしてベンチマーキングの行政部門評価における優位性や協働との関係性について確認する。次に行政ベンチマーキングでは多様な評価主体の参画が想定されるため、各評価主体や評価性向を導出する理論的枠組をもとに行政部門評価としての協働の有意性について検討する。同時に行政部門評価におけ

66 西出 順郎

る協働の実効性について検討する。最後に当該検 討を総括し、行政ベンチマーキングにおける協働 の内実化に資する実務的示唆について論及する。

なお本稿でいう「評価」は概念的には評価論の一般的定義が前提となるが、実務的文脈においてはわが国地方自治体の基本的枠組を想定した「行政部門の評価(もしくは行政部門評価)」と相互互換的に取り扱う。「ベンチマーキング」と「行政ベンチマーキング」については前者を経営学上の概念的定義、後者を行政部門評価に応用された実体的定義として区別する。よって行政ベンチマーキングはあくまで当該評価の一形態として取り扱う。

## 2 行政部門の評価とベンチマーキング

#### (1) 行政部門とベンチマーキング

ベンチマーキングとは別の組織の優良事例と比 較分析し自らの取組改善を試みる組織経営の技法 である。わが国の行政部門においては評価研究の 先進国たる米国と同様、形成的役割を担う評価技 法のひとつとして着目されてきた。1990年代後半 には米国オレゴン州の「オレゴン・ベンチマーク ス」と称する「アウトカム指標を測定・評価する 試み」が注目される。このことから「アウトカム 志向の業績測定 | イコール「ベンチマーキング | といった解釈が一般的となった。たとえば西尾 (2000:24) では「市民参加の下に策定された自 治体計画で行政サービスの達成目標として設定さ れた指標」で測定する評価取組がベンチマーキン グとして紹介されている。2000年代前半には米国 の職能団体であるICMA(International City/County Management Association) やオレゴン州内の各カ ウンティが測定結果の相互比較を行うようにな り、わが国では総合研究開発機構を中心とした 研究2や都市自治体間での取組実践が積極的に推 し進められた。よって行政ベンチマーキングは 「数値によるアウトカム測定」から本来の定義で ある「他者との比較による取組改善 | へと自らの 役割を変容させ、わが国においても普及、発展す ることになる。

#### (2) 業績測定とベンチマーキング

現行の行政部門評価は目標値と実績値を比較する業績測定という技法をもとに、測定結果を自らの経験知でもって検証する枠組である。よって業績測定からベンチマーキングへと評価枠組の外延が拡大したのは世界における「共通の進化"Common Evolution"」(Bouckaert 1996:234)のひとつと考えられよう。では拡大といった動きを推し進めるベンチーキングの特性とは何か。それは現行評価の実務的問題を低減させ得る評価設計にある。具体的には次の3つが想定される。

第一にベンチマーキングの数値目標は政府目標 という位置付けにはならず、またその達成の可否 で行政効果の是非をも問えない点である。測定結 果は最良事例を発掘するためのモニタリング情報 に過ぎない。したがって「政府全体もしくは首長 の政策的目標値なのか行政部門の技術的目標値な のか」といった政策策定プロセスの曖昧さによる 目標値の定義問題や目標達成の是非にかかる責任 問題を回避できる。

次に最良事例を抽出するため深堀調査を実施する点である。ベンチマーキングは評価客体の詳細な分析情報を必要とする。よってアンケート調査、フォーカス・グループ・インタビューといった社会調査手法やロジック・モデリングによるプログラム・セオリーの論理的検証が不可欠である。現行評価の多くが業績測定と経験知に基づく直感的判断に依存するなか、ベンチマーキングではその非科学性を超える体系化された評価技法が必然的に遂行される。

最後に最良事例のエビデンスに依拠する改善型の評価結果が産出される点である。これによって 業績を重視する現行評価にはないフォワード・ルッキングの具体的改善策が提示され、評価結果の 実用性は向上する。

以上の実務的問題、すなわち「評価目標の正当性」「評価技術の向上」そして「評価結果の有効活用」は行政測定研究が台頭した1950年代から議論の対象となっており、未だ抜本的な解決には至っていない。その意味で評価の外延拡大はベンチマーキングの特性を活かして測定問題の解消を目論んだ次なる展開であったといえよう。

#### (3) 行政ベンチマーキングと協働

ベンチマーキングの特性が現行評価の問題解消 に貢献するとしても、行政ベンチマーキングがそ う容易く現実機能するわけではない。なぜなら自 己評価が基本の現行評価では想定し得ない「複数 の評価主体による協働」が当該現実機能の大前提 だからである。深堀調査等の評価作業を複数の主 体が連携すること自体容易なことでない。特に有 効性や効率性より無謬性に重きを置く行政の本質 に鑑みると、協働による評価作業にはそれ相応の 相互理解や信頼関係が求められよう。さらには昨 今の行政の不祥事問題を端緒に「市民目線の評 価 |、すなわち市民公募型の評価枠組を社会は強 く要請している。連携行政機関のみならず利害の 異なる外部評価主体との協働を受容することは行 政ベンチマーキングにおける論理的帰結でもあろ う。よって「行政部門の評価として評価主体の協 働は有意ある行為なのか。|、この問いかけこそが 当該現実機能の基本的命題となるのである。

### 3 協働の有意性 - 評価性向の視点から -

#### (1) 多様な評価主体

行政ベンチマーキング研究において、政治的任 用者や予算担当者さらには顧客など多様な利害関 係者間の連携は当該成否にかかる重要な要因 (Keehley et al. 1997) とされるものの、多様な評 価主体の協働に着目した文献は管見の限り見当た らない。その一方、概念的ながら行政部門評価と しては行政外部による評価の必要性が終戦直後か ら指摘されていた。たとえば、吉富(1951)では 行政活動の民主性を確保すべく素人による評価 が、小林(1974)では市民参加による測定指標の 策定が、そして西尾(1976)では行政ニーズから 政策目標の転換過程での市民参画が、それぞれ行 政統制という規範的文脈から要請されている。 1990年代後半には行政評価・政策評価という名の もと行政部門の評価システムが全国に普及し、評 価における市民という概念もより深化した。たと えば西尾 (2000:29-32) では、「従来の市民参 加・住民参加にとらわれず | メディア、NPO等を 含めた「民間のすべての主体」を市民の構成要素

とし、「納税者」「消費者」などに類型化された市 民が評価客体との関係性から論じられている。

翻って対人的行政サービスの評価を研究対象と して米国で発展を遂げた評価論では、評価主体の 前提がプロフェッションたる評価専門家であるこ とから当該専門家と利害関係者との協働が研究課 題のひとつとなっている。参加型評価、エンパワ メント評価など概念的評価技法の発展は当該研究 の成果といえよう。もっとも評価論から評価主体 の存在を掘り起こすことも決して難しいことでは ない。評価主体は原則「評価の結果を欲する者」 に他ならないからである。たとえば評価論の大家 であるWeiss (1998) は評価結果の想定利用者と して、国および地方の行政機関、当該政策の立案 者や評価客体の実施者、当該分野の研究者さらに は一般市民など、多様な利害アクターを列挙して いる。行政部門の内外に多くの評価主体が潜在す る一証左といえよう。

#### (2) 評価主体の理論的枠組

評価とは組織の取組業績を検証することであり、その正当性は検証目的を導出する理論的枠組によって付与される。よって業績測定やベンチマーキングといった評価技法にかかわらず、「行政部門の取組」に着眼することで経営学や行政学等から理論的枠組が導出され、それらを帰納的に考察することで各評価主体の全体像が明らかにされる。

評価が依拠する理論的枠組の導出については、制度的特徴から国立大学法人評価のそれを明らかにした研究(西出 2009)がある。当該アプローチを現行の行政部門評価に援用するなら主たる特徴として次の3つが抽出されよう。すなわち「行政は政治の意思決定に基づき執行を委ねられた部門であること」、「当該執行部門の長たる首長は公選職であること」、そして「予算編成権や委任条例等に基づく裁量権を首長が有していること」である。これらの特徴が評価の制度化を要請し、当該根拠として導出される理論的枠組が「管理過程論」、「エージェンシー理論」、「行政統制論」である。

#### ①管理過程論

管理過程論は経営学、特に経営管理における伝

68 西出 順郎

統的理論のひとつである。ここでは組織管理者の管理職能は「計画」「組織」「指揮」「調整」「統制」の5つの機能的要素から構成されるとし、これらの機能を連続的に循環する重要性が説かれている(今田 2009)。いわばPlan-Do-See、Plan-Do-Check-Actionといったマネジメント・サイクルの原型的発想といえよう。当該論における評価は計画の実効性を検証する「統制」機能の一手段として位置付けられる。想定される評価主体としては、行政部門の最終責任者たる「首長」、評価客体の現場責任者たる「事業部門の長」、さらには行政ベンチマーキングに参画する「連携行政機関の事業部門長」が挙げられよう。

### ②エージェンシー理論

エージェンシー理論では依頼人と代理人との契 約関係から生じる非効率現象を経済学のアプロー チから検討する。当該理論は、契約関係にある依 頼人と代理人との利害が必ずしも一致せずして前 者が大幅な裁量権を後者に委ねた場合、両者間に 情報の非対称性が生じることを指摘する。よって 代理人が契約期間中に自己利益を悪徳的に追求 し、依頼人の利害を軽視するといったモラル・ハ ザードを抑制するため、モニタリングという評価 手段が検討される (Picot et al. 2005=2007)。 行政 部門の評価では次のような依頼人 - 代理人関係が 想定されよう。(a) 「議会 | - 首長、(b) 「首長 | -部下たる各部門の長、(c)「総務部門」- 事業部門、 (d) 補助金交付の「上級行政機関」- 首長、そし て(e) 「主権者 | - 首長における法的もしくは規 範的関係である。依頼人たる「評価主体 | は、(a) では「議会の決定に従って首長が行動したか」、 (b) では「首長の指示に従って各部門の長が行動 したか」、(c) では「事業部門の過年度取組に無 駄や問題がなかったか |、(d) では「規則に従っ て首長が補助金を執行したか」、そして(e)では 「主権者の意向に沿って首長が行動したか | とい った総括的な視点から評価客体を検証する。

#### ③行政統制論

行政統制論は行政部門の組織特性から導出される行政学上の理論である。行政統制とは「行政活動を他律的に抑制する仕組み」(桑原 2007:288)である。すなわち国民と政府との信託関係を良好に維持すべく行政が負うべき責任を国民自らが統

制する手段である。現代の行政活動は多様化、複雑化そして高度化しており、国民の代表者たる議会による統制にも限界がある。よって行政統制は行政国家において民主主義を保障する重要な手段とされている。当該論はエージェンシー理論の応用モデル(今村 1999)としても論じられ、「主権者」として自らが選んだ首長の監視が基本型となる。しかし行政の監視により着眼するならば、選挙による直接的な依頼人・代理人関係がない当該行政地域外の「市民活動団体」や「報道機関」であっても評価主体として十分想定されよう。

## ④市民参加論・NPM論・消費者行動論・ジャー ナリズム論

このように3つの理論的枠組によって想定される評価主体の具体像はみえてきた。但し「主権者」については西尾(2000)のようなさらなる類型化が可能であり、各々の存在意義は「有権者」に代表されるエージェンシー理論もしくは行政統制論を基本としつつも次の規範的もしくは実証的論拠、すなわち「市民参加論」「NPM論」「消費者行動論」「ジャーナリズム論」によって補強される。

市民参加論における「主権者」には「地域社会の一員としての市民自身の責務」や「協働による公共サービス」(佐藤 2006:9)を遂行する立場としての一面がある。前者としては市民自治意識の確認・向上を目的とした「地域生活者」が、後者としては評価客体の実施者として事業型NPOなどの「市民活動団体」が挙げられる。

NPM論ではその行動規範である「顧客主義」 (大住 2002:12) や「VFM (Value for Money)」 (*i bid*:14) への反射的行為として「納税者」および「受益者」が存在する。特に後者の場合は消費者行動論でいう「購買後行動」(冨貴島 2006:18) や足による投票の代替となるクレーム的行為の一手段として評価への参画が想定される。

公権力への監視という点では、「報道機関」が ジャーナリズム論の「公権力との対峙、社会的責 任」(駒村 2006:47-48)によって、また提言型 の「市民活動団体」が市民参加論の「行政及び議 会に対する統制」(佐藤 2006:8)によってさら に支援されよう。

## ⑤シグナリング理論

「首長」においては自らの実績を有権者へ積極的に知らしめるため自ら評価を行う場合がある。この積極姿勢は行政統制論ではなく経済学でいう「シグナリング理論」に依拠する。すなわち情報の非対称性の緩和を欲する情報保有者が自らの優位性を周知するために自らの情報の開示となる行動を先に採る(Picot et al. 2005=2007)手段としての自己評価である。たとえば企業のIR情報の公表はシグナリングであり、それによって企業は自らの質に見合った投資を得ることができる。「首長」の場合、自己評価の公表がシグナリングであり、選挙時の票の獲得がその成果といえよう。

### 6評価論

専門職としての外部評価者の必要性を論じるのが「評価論」である。当該論では外部評価によって評価結果の「独立性、信頼性、客観性が維持できる」(Patton 1997:138)とされ、内部評価の問題解消を目的に評価や専門領域の研究者といった「評価専門家」の存在を積極的に認めている。

表1は想定される評価主体とその理論的枠組について整理したものである。

## (3) 評価主体と評価性向

Rossi et al. (2004: 80) の「評価階層

(Evaluation Hierarchy)」から解釈すると、評価論における評価とは評価客体の「必要性」「論理性」「遂行性」「有効性」「効率性」、そしてこれら5つの全体総括としての「妥当性」を検証するものである。しかしながら評価結果を書き綴った「ドキュメント」は当該評価基準にのみ傾注すればよいわけではない。各評価主体が望むドキュメント全体の有様、すなわち評価結果の指向性についても十分留意しなければならない。それら指向性への誠意ある対応なしに評価結果の客観性や信頼性、さらには実用性を担保することが極めて困難だからである。このように評価主体が期待する評価結果の指向性を評価性向と呼ぶ。

各評価主体は自らの理論的枠組を背景に評価作業へ参画するため、単体のドキュメントであっても多様な評価性向が要請される。もちろんその時々の社会的もしくは政治的文脈にも影響されるが、各評価主体が要請する評価性向は各理論的枠組によって概ね次のように導出されよう。

## ①管理過程論

管理過程論では行政活動の自己改善に資する評価情報が必要となる。よって「妥当性」に対する 実務的専門性(以下、「実務性」)、それに基づく 「形成性」、時として戦略としての「秘匿性」が求められる。主として評価客体の執行責任者たる首

耒1	相定され	ス評価主休と	その理論的枠組

社会的属性	評価主体	評価客体との関係	理論的枠組
立法部門	議 会	案の法的承認	エージェンシー理論
	首 長	案の決定・提出	管理過程論・エージェンシー理論・シグナリング理論
行政部門	総務部門	案の資源配分担当	エージェンシー理論
	事業部門	案の設計実施	管理過程論
外部の	連携行政機関	比較結果の活用	管理過程論
行政部門	上級行政機関	資金の補助	エージェンシー理論
) I = +-	有 権 者	意思決定者の選択	エージェンシー理論・行政統制論
主権者	納 税 者	資金の提供	NPM論・エージェンシー理論・行政統制論
としての	受 益 者	サービスの消費	NPM論・消費者行動論・エージェンシー理論・行政統制論
外部部門	地域生活者	自治意識の向上	市民参加論・エージェンシー理論・行政統制論
	市民活動団体	協働主体、権力監視	市民参加論・行政統制論
	報道機関	権力監視	ジャーナリズム論・行政統制論
専門職として	評価・当該政策	評価分析支援	評価論
の外部部門	の専門家	計画力	百十

(注) 社会的属性における斜線は上下両属性もしくはいずれかの属性に位置する当該主体があることを意味する。

(出所) 筆者作成

70 西出 順郎

長や現場責任者たる事業部門の長が業務改善のために要請する評価性向である。他の組織の最良事例を導入したい現場責任者においても同様ではあるが、この場合、評価客体の「遂行性」に特に注目するであろう。

## ②エージェンシー理論

エージェンシー理論は代理人の目標達成度や取 組の「妥当性」を判断するため評価結果の「総括 性 | を追求する。さらに依頼人と代理人との関係 性によっては管理過程論や後述する行政統制論の 影響を受けた評価性向が要請される。たとえば有 権者もしくは議会による首長に対する評価では 「平易性」と「公開性」が期待される。逆に上級 行政機関による首長に対する評価、首長による事 業部門や総務部門の長に対する評価、そして総務 部門による事業部門に対する評価では「実務性| を帯びた情報を求めるとともに戦略的意図がない 限り「公開性」に対して関心を払わない。また当 該4つの関係性において依頼人は代理人に対して 「誤謬性」を想定しながら評価を行い、特に総務 部門による事業部門の評価では次年度計画を見越 して評価客体の「必要性」により着目する。

### ③行政統制論・NPM論および消費者行動論

行政統制論は外部による首長の監視が目的であるため、評価結果には「総括性」や「平易性」、「公開性」が要請される。監視の基本基準は評価客体の「妥当性」であるが、首長の独断専行的行為や「無謬性」による評価結果の誘導を警戒し、「公正性」や「誤謬性」の視点からも評価する。当該論の評価主体は主権者であるが、なかでも業績に着眼する納税者はNPM論によって「効率性」を、受益者はNPM論と消費者行動論によって「有効性」をより指向する。但し後者は自己効用の増大を志向するため、当該年度の結果以上に自らの便益が高まる次年度サービスの向上、すなわち「総括性」よりも「形成性」により関心が傾く。

## ④シグナリング理論

シグナリング理論では首長が自らの行政活動の「公正性」および「妥当性」を標榜するために評価結果を利用する。よって「総括性」「平易性」「公開性」が前提となる。その一方、首長自らが評価主体となるため、評価の基本的価値前提ともいえる「誤謬性」の可視化に対する懐柔行為、す

なわち「無謬性」への暗黙的誘導がなされる可能 性がある。

### ⑤ジャーナリズム論・市民参加論

ジャーナリズム論は行政統制論と同様、悪徳的なシグナリングの監視を目的としており、その判断基準は「妥当性」のみならず「公正性」、さらには市民目線を強く意識した「総括性」「公開性」「平易性」そして「誤謬性」を指向する。市民参加論においてもジャーナリズム論と同じくシグナリング理論の弊害に対する防あつ機能が求められている。但し、地域生活者は自己啓発が主たる参画目的であるため自らの評価性向を積極的に標榜することはなく、自分の自治意識を満足させる「知識創出型」とは、評価結果の具体的活用を求める「知識創出型」とは、評価結果の具体的活用を求める「総括性」や「形成性」と異なり、評価による知識の蓄積によって学びや啓発を促進する観念的活用を企図する評価性向を意味する。。

#### 6評価論

評価論による外部評価主体には、評価結果の「中立性」の確保や科学的専門性(以下、「科学性」)の向上が要求される。但し、評価作業の発注者の存在が前提となるため、当該評価主体の評価性向は当該発注者との権力的関係に強い影響を受ける。たとえば「実務性」か「平易性」、「総括性」か「形成性」、「秘匿性」か「公開性」といった評価性向の選択は発注者の意向に大きく依存することになる。特に評価作業の受注者側が厳しい業務競争に晒されている状態もしくは発注者側が革務局機能を独占している状態では、発注者側の首長がシグナリングを強く嗜好することで「中立性」や「科学性」といった評価性向の質が作為的に変容することも想定される。

### (4) 評価性向の類型化

各理論的枠組によって多様な評価性向が導出された。これらは類似もしくは対立する性質を有しており、分類すると次の5つの範疇、「基準型」、「活用型」、「接続型」、「表現型」、「誘導型」に整理することができる。

#### ①基準型

基準型は「評価客体をどのような意図をもって 検証するのか」、その物差しを具体的に提示した 評価性向の類である。必要性や取組状況(遂行性)、 具体的成果(有効性)や費用便益(効率性)など の評価基準が挙げられる。さらには全体としての 妥当性、公的部門としての公正性が想定される。 ②活用型

活用型は評価結果の使い道に依拠するものである。取組価値の是非を問う総括性と取組の質の向上を支援する形成性、さらには評価主体の自己啓発や評価客体の情報蓄積を目的とする知識創出性が挙げられる。

### ③接続型

接続型は評価結果に対する部外者からのアクセスビリティに関するものである。評価主体が評価結果に対して部外からの接続を欲さない場合には秘匿性が、逆に積極的に接続性を開放して社会全体に何らかのメッセージを伝えたいもしくは説明責任を着実に履行させたい場合には公開性が該当する。

#### 4)表現型

表現型は評価結果のドキュメントの記述レベルに関するものである。評価主体が一般市民であれば誰でも納得し得るための平易性、行政部門であれば具体的改善に資するための実務性、さらには信頼性や客観性確保のための研究者による科学性

が挙げられる。

### ⑤誘導型

誘導型は評価主体が評価結果の操作誘導もしく はその抑制を目論むことを表す評価性向の類であ る。ここでは首長自らが取組努力を誇張してその 成果を主張する無謬性や議会もしくは主権者など の評価主体が首長の瑕疵を厳しく検証する誤謬性 が該当する。

表2は各評価主体が期待する評価性向を範疇別に整理したものである。なお首長の場合には3つの理論的枠組が意図の異なる評価目的を各々導出することから枠組ごとに区別した。(a) が管理過程論、(b) がエージェンシー理論、(c) がシグナリング理論に依拠した評価性向である。

## (5) 促進的協働と調整的協働

各評価主体および評価性向において多様性がみられる以上、ひとつの評価客体を協働で評価するのは決して容易なことではない。但し評価主体の有機的な協働は全く不可能というわけでなく、ある状況下においては論理的に受容される。具体的には次の2つである。第一に同じ評価性向を期待する評価主体が協働する場合である。同一評価性向を志向する協働作業では、評価結果の質の向上、

評価主体	基準型		活用型	接続型	表現型	誘導型
議会	妥当性		総括性	公開性	平易性	誤謬性
	(a)	妥当性	形成性	秘匿性	実務性	_
首 長	(b)	妥当性	総括性	_	実務性	誤謬性
	(c)	公正性・妥当性	総括性	公開性	平易性	無謬性
総務部門	妥	当性 (特に必要性)	総括性	_	実務性	誤謬性
事業部門		妥当性	形成性	秘匿性	実務性	_
連携行政機関	妥当性(特に遂行性)		形成性	秘匿性	実務性	_
<b>上級行政機関</b> 妥当性		総括性	_	実務性	誤謬性	
有 権 者	公正性・妥当性		総括性	公開性	平易性	誤謬性
納 税 者	公正性・妥当性(特に効率性)		総括性	公開性	平易性	誤謬性
受 益 者	公正性・妥当性(特に有効性)		形成性	公開性	平易性	誤謬性
地域生活者		公正性・妥当性	知識創出性	_	_	_
市民活動団体	市民活動団体 公正性・妥当性		総括性	公開性	平易性	誤謬性
報道機関		公正性・妥当性	総括性	公開性	平易性	誤謬性
当該政策もしくは	*		*	*	科学性	中立性
評価の専門家						

表2 評価主体が期待する評価性向

(出所) 筆者作成

<sup>(</sup>注)「一」は積極的には特定の指向性を持たないこと、「\*」は発注者の意向に依存することを意味する。

72 西出 順郎

特に情報収集・解析の強化が期待できる。たとえば形成性を期待する行政部門とそのサービス受益者との協働では、需給両者の視点に立った評価結果の産出が促進される。

第二に対立する評価性向をもつ評価主体が協働する場合である。相互の牽制的思考がもたらす調整作用によって評価性向の偏重が抑制される。また使い道としての汎用性も確保される。特に内外評価主体による協働では、外部評価の弱点である内部情報の収集が容易となり、専門性の高い情報が平易な情報へと変換される。さらには無謬性による瑕疵を発見する協働作業が期待できる。

したがって、評価性向の性質を十分見極めて、かつそれを有機的に活用する意図も持って協働するのであるなら、行政部門評価の質の向上を目論むうえで協働は有意的行為であると判断できよう。

## 4 協働の内実化を妨げる実施上の脅威

### (1) 3つの脅威

多様な評価主体による協働の有意性は論理的には確認できた。しかしそのこと自体が協働の実効性を直ちに担保するというわけではない。なぜなら本来、評価には組織の常態に抗う性質が存しており組織の中に息づく場所などない(田辺 2006)。よって組織の論理が評価の論理を統制する、するわち前者が後者を凌駕して優位に働く状態が日常的に起こり得るからである。特に複数の理論の理品が政治・行政部門の評価が構築される以上、評価の理屈が政治・行政的文脈からの何らかの関与によって評価の途中で骨抜きにされることも常いによって協働の内実化を対ける実施上の脅威は行政部門の評価作業に内在するのか。もちろん個別具体の組織的状況にもよるが、少なくとも次の阻害的事象が想定されよう。

### ①非公式な力による操作誘導

評価主体間の非公式な力関係に依拠する脅威である。行政内部の組織構造上、総務部門と事業部門の間には直接的指揮命令関係は存在しない。外形的にも促進的・調整的協働を阻害しあう関係にはない。しかし両者間には予算・人事といった経営資源の配分をめぐる間接的権力関係が常に介在

する。よって形成性に着眼する評価であっても経 営資源の差配に影響がある場合、対等な立場での 協働は期待できない。さらに行政外部との協働に おいても、外部評価主体の事務局機能を行政部門 が直接担う場合には外部主体の意向どおりにこと は運ばない。時には当該事務局による情報操作、 評価結果に関する意図的誘導が起こり得ないとも 限らない。

## ②非合目的な外部評価主体の選択

外部評価主体の選択過程に内在する脅威である。役職で外部評価主体を委嘱する外形的なスクリーニングでは議論が批判的かつ能動的に展開されるか否か疑問である。市民参画を標榜して外部主体を自由に公募する場合であっても、任意の主権者に地域生活者以上の活躍を期待することは困難である。また選択過程で当該事務局の意図的な選択が実行され得ることも否定できない。このように非合目的な外部評価主体の選択がなされた場合、結果的に外部評価主体の脆弱な潜在的論争力が問われることになる。

### ③協働作業に対する結果責任の所在

評価作業の結果責任に関する脅威である。評価 作業は改善措置を行政部門に強制できるものでは なく、評価作業の「アウトカム」である評価結果 の「有効活用」は首長が事業計画を改善すること で具現化される。また外部の評価主体は当該アウ トカムを具現化する責務からそもそも開放されて いる。よって首長や行政内部の評価主体は「評価 結果の予算への反映状況 | を積極的に公表し、評 価作業自体の実績強調に腐心する。その結果、当 該反映状況の作成段階で予算査定作業の「結果 | を包摂し、評価結果に予算査定の結果を「反映 | させ、可逆的に外形的整合性を確保して当該アウ トカムの達成を図る。かくして次年度予算の「査 定結果」と前年度取組の「評価結果」の予定調和 が成立する。協働作業による評価結果に対して外 部評価主体の責任は求められることなく、そして 評価主体間の対等かつ責任ある関係も虚構のまま 成熟することなく、協働による評価作業は緊張感 のない儀式化された遊戯へと化すことになる。

## (2) 脅威の根本原因

当該脅威は現行の行政部門評価においてかなり

の確からしさで起こり得る事象である。よって協 働の実効性を高めるには当該事象の発現を事前に 抑制する手立てを講じるべきであり、そのために は当該脅威の根本原因を見極める必要がある。そ こで各阻害的事象のメカニズムを垣間見るとそれ ら背後に通底する行動規範から当該原因の輪郭が 浮かび上がってくる。それは権威に抗う明示的行 為に対する暗黙的抗いであり、その源は最終的に は自らの組織に対する忠誠であり、そのように誘 導するのが行政部門評価の実質的所有者たる首長 であるという現実である。行政の内部部門は首長 の部下である以上、最終的には評価が首長のシグ ナリングであることを意識する。特に総務部門は 政治・行政的文脈を背景に事業部門を誘導し、結 果として無謬性を帯びた評価結果の「公表」を試 みる。「公表」が前提となる評価作業においては、 総務部門は誤謬性をもって事業部門を評価する 「依頼人」の立場ではなく、シグナリングとして の評価結果の公表を目論む首長の「官房部門」と しての立ち回りを演ずるからである。もちろん外 部評価主体との協働においても行政内部全体が同 様の行動規範をもって臨む。よって評価設計にお いて協働の機能性を論理的に具備し得たとして も、行政部門の評価、特に実際の作業過程のなか で協働の内実化を維持し続けるのは厳しいといわ ざるを得ないのである。

## 5 行政ベンチマーキングの現実機能 - 結 びにかえて -

以上を総括すると次のとおりである。行政ベンチマーキングは行政部門評価の外延を拡大する評価枠組である。しかしその現実機能を向上させるには、行政部門の評価に介在し得る多様な評価主体が果たして有機的に協働し得るのか、すなわち協働の実効性を十分に吟味しなければならない。行政部門評価の存在意義は管理過程論やエージェンシー理論、行政統制論など意図の異なる複数の理論的枠組から改めて確認された。また、当該各枠組によって首長から納税者、地域生活者といった多様な評価主体が明示され、総括性や公開性、誤謬性といった各評価主体の評価性向が導出され

た。そして当該評価性向を論理的に検討すること で促進的および調整的な協働においてのその有意 性が明らかとなった。しかし同時に行政部門の評 価であるが故に協働の内実化を妨げる評価実施上 の脅威の存在さらにはその根本原因も浮き彫りと なった。

複数の理論的枠組を根拠とする行政部門の評価において協働の内実化を担保するのは決して容易なことではない。そしてそのためには協働の有様や評価枠組の設計等を個別具体に見直すことが要請されるだろう。しかしその一方、実施上の脅威が現行の評価枠組を前提としていること、また行政ベンチマーキングが現行評価とは異なる新たな枠組であることに今一度着目すると、当該脅威の低減可能性は行政ベンチマーキングであるが故に期待できないわけではない。なぜなら行政ベンチマーキングは前述したその特性に鑑みてもシグナリングや目標管理ではない取組改善を志向する評価作業に特化することが可能であり、また、その発想こそが解決の糸口といえるからである。

結びにかえて、最後に行政ベンチマーキングの 現実機能、すなわち協働の内実化を担保するには どのような評価作業を試みるべきか、その実務的 示唆について論及したい。

行政ベンチマーキングの目的は連携機関等との協働により最良事例を抽出し、自己の改善を図ることである。事業部門の査定が目的ではない。よって行政ベンチマーキングでは、無謬性や総括性といった評価性向を組織的に低減させる、特に総務部門には形成性を意識させる協働を強く要請すべきであり、予算部門との意識共有が最良事例を具現化させるための重要要因となる。

また行政部門の民主的統制が目的でもない。よって首長の監視といった外部からの誤謬的視点を低減させ、受益者や地域生活者の視点に立った改善志向の主権者との協働を展開すべきである。このことは当該主権者の自己効用の増大にも貢献し、Win-Winの評価作業が実現できよう。

さらには首長の成績採点のために実施されるものでもない。よって連携機関等の合意に基づく評価作業であることを強調し、いかなる首長のオーナーシップも積極的に排除すべきである。これによって単一組織による誘導型の消費性向は消滅

74 西出 順郎

し、当該分野の専門家や研究者の協力獲得も容易 となる。

このように行政ベンチマーキングでは評価主体間の評価性向を把握し、その文脈における有機的な仕組を講じることで協働の実効性はかなり向上しよう。行政ベンチマーキングだからこそ評価主体の協働が要請され、同時に行政部門評価の協働の難しさが低減される。この関係性こそが行政部門評価の外延拡大をもたらした真の理由かもしれない。

行政ベンチマーキングのさらなる試行錯誤は行 政部門評価のまさしく新たな突破口といえよう。

### 謝辞

本稿の草稿に対しては匿名の査読者の方々から 大変貴重なコメントを頂いた。この場を借りて感 謝申し上げる。

### 注記

- 1 詳しくはAmmons (2001:2-3) を参照されたい。
- 2 当該研究成果の一部は2003年発行NIRA『政策研究』 通巻190号の特集として収録されている。
- 3 直近の2011年発行『日本評価学会』第11巻第1号に 「市民社会におけるNPOと評価の展望」と題した特 集が組まれた。行政ベンチマーキングや評価主体の 評価性向に着目するものではないが、参加型の評価 に関する論文が複数収録されている。
- 4 本稿では社会的もしくは政治的文脈を重視する評価 の定義的特性 (Rossi et al. 2004) に鑑みて、研究者 の学術研究活動は評価活動の範疇外として取り扱 う。
- 5 当該評価性向についてはPatton (1997:70) における "Knowledge-Oriented Evaluation"の概念を参考にしている。

### 参考文献

今田治(2009)「マネジメント論の発展:古典的理論の生成」、田中照純・小久保みどり編『マネジメント論ー管理、そして組織を考えるー』、ナカニシヤ出版、21-27

- 今村都南雄(1999)「行政責任と統制」、今村都南雄・ 武藤博巳・真山達志・武智秀之『ホーンブック行政 学』、北樹出版、204-233
- 大住荘四郎(2002)『パブリック・マネジメント』、日本評論社
- 桑原秀明(2007)「行政統制:行政統制と行政責任」、 『政治学・行政学の基礎知識 第2版』、一藝社、 288-289
- 小林紘 (1974)「自治行政とシステム分析」、『地方自治』、317:22-27
- 駒村圭吾(2006)「法制度から見た「ジャーナリズムと権力」」、大石裕編『ジャーナリズムと権力』、世界思想社、37-59
- 佐藤徹 (2006)「市民参加の基本的視座」、原田寛明監修・佐藤徹編集代表『地域政策と市民参加』、ぎょうせい、3-12
- 総合研究開発機構編(2003)『NIRA政策研究』、190
- 田辺国昭 (2006)「政策評価制度の運用実態とその影響」、『レヴァイアサン』、38:86-109
- 冨貴島明(2006)「消費者行動論(1)消費とは何か」、 『城西大学経済経営紀要』、24:17-37
- 西尾勝 (1976)「行政需要概念の再構成」、『年報行政 研究』、12:1-51
- 西尾勝 (2000)「行政の評価方式の拡張をめざして」、 西尾勝編著『行政評価の潮流-参加型評価システム の可能性-』、行政管理研究センター、1-38
- 西出順郎 (2009)「国立大学法人法化制度の理論的考察-制度設計上の合目的性と機能可能性-」、『日本評価研究』、9(3):95-108
- 吉冨重夫(1951)「行政能率と行政計画」、『都市問題 研究』、21:17-44
- Ammons, N.D. (2001). *Municipal Benchmarks*, second edition, Thousand Oaks: Sage Publication.
- Bouckaert, G. (1996). Measurement of Public Sector Performance: Some European Perspectives, Halachmi, A. and Bouckaert, G., eds., *Organizational Performance and Measurement in the Public Sector*, Westport: Quorum Books, 239-273.
- Keehley, P., Medlin, S., MacBride, S., and Longmire, L. (1997). *Benchmarking for Best Practices in The Public Sector*, San Francisco: Jossey-Bass Publishers.
- Patton, M.Q. (1997). *Utilization-Focused Evaluation*, third edition, Thousand Oaks: Sage Publication. (=大森彌

監修、山本泰・長尾眞文編(2001)『実用重視の評価入門』、清水弘文堂)

Picot, A., Dietl, H., and Frank, E. (2005). 丹後安治ほか 訳 (2007)『新制度派経済学による組織入門 第4版』、 白桃書房

Rossi, P. H., Lipsey, M.W., and Freeman, H.E. (2004).

Evaluation: A systematic Approach, Seventh edition, Thousand Oaks: Sage Publication. (=大島巌ほか監訳 (2005) 『プログラム評価の理論と方法』、日本評論社) Weiss, H.C. (1998) . Evaluation, second edition, Upper Saddle River: Prentice Hall.

(2011.9.2受理)

# The Functional Capability of Benchmarking in the Evaluation of Public Administration — Collaboration and Evaluation Propensity in Self-interest Based Evaluators —

Junro Nishide

Iwate Prefectural University nishide@iwate-pu.ac.jp

### Abstract

The purpose of this research is to analyze the theories that derive self-interest based evaluators in the benchmarking of public administration, and to review the functional capability of benchmarking in the evaluation of public administration.

As a result, it was found that the collaboration of self-interest based evaluators in the evaluation of public administration was significant to acceleration and coordination from the viewpoint of the evaluation purpose and propensity. Moreover, it was recognized that threats to the evaluation implementation existed. This was caused by the head of the administration for the evaluation of public administration's ownership, which impeded the functions of the collaboration. Finally, the possibility of eliminating the threats to the evaluation and enhancing the utilization of benchmarking in public administration are considered.

## **Keywords**

Benchmarking, Evaluation, Collaboration, Self-interest Based Evaluator, Evaluation Propensity

## 【研究論文】

## 行政評価過程における職場討議の導入とその効果検証

佐藤 徹 高崎経済大学 tsato@tcue.ac.jp

## 要約

わが国では1990年代半ば以降今日に至るまで、多くの地方自治体において行政評価システムが急速に導入されてきた。この間、評価の制度的導入という点では一定の定着が見られたが、事業担当課職員の間には評価制度に対する「やらされ感」や「負担感」が依然として根強いといった深刻な問題を抱えている。そこで、近年筆者はこうした実践的課題に対処するため、創造型政策評価(Creative Policy Evaluation:CPE)を提案している。本研究ではCPEを現実の行政過程に導入し、その効果を職員へのアンケート調査によって検証した。その結果、評価目的の理解、論理的思考方法や評価指標の導出方法の習得、評価における職場議論の方法の習得という点で統計的に有意な差が見られた。さらに評価過程における職場議論が活発であるかどうかがどのような因子によって決定されるかについても分析したところ、ファシリテーターが議論の場を適切にマネジメントしたかどうかに最も大きく依存していることが明らかとなった。

## キーワード

行政評価、職場討議、評価ワークショップ、創造型政策評価

## 1. 研究の背景と目的

わが国では1990年代半ば以降今日に至るまで、多くの地方自治体において行政評価システムが導入された。その結果、「休廃止事業の選択・予算編成への活用」「住民や議会に対する説明責任」「総合計画の進行管理」や「職員の意識の向上」などに一定の成果を挙げたとする調査報告もある(松尾2009、田渕2010)。

しかし一方で、行政評価の成果に否定的な見解 も一定割合存在しており、特に行政評価の目的と して重視されていた意識改革に関連するものとし て、住民の立場に立った企画や対応、仕事に対す るモチベーション、部局間の垣根を越えたコミュニケーションなどについて十分な成果があったと実感できるには至っていないとの指摘がある(松尾 2009)。たしかに、「評価制度を導入したが、事業改善にはあまり結びついていない」「余計な事務仕事が増えた」「職場内での議論がほとんどなく、一部職員が評価シートを作成しているにすぎない」「評価シートが評価担当課からの照会文書と化している」など、現場職員の評価制度に対する「評価」は必ずしも芳しいものではない。

さらに田渕(2010)によれば、行政評価を導入 して約10年が経過し、評価制度を見直し中とする 自治体の担当者からは「行政評価に対する理解や 78 佐藤 徹

意識が低いため、職員に "やらされ感"がある」「作業の負担感が大きい」などの意見が寄せられているという。そのため評価に伴う事務量を軽減するため評価シートの枚数や評価対象の事業数を削減するなど、一部では評価の簡略化傾向も散見される。だがそれがますます評価のルーティン化やマンネリ化に拍車をかけているように思われる。。

こうした閉塞的状況から脱却する方策の1つとして、筆者は近年「創造型政策評価(Creative Policy Evaluation: CPE)」を提案し、自治体職員向けの研修やセミナーなどを通じて、その実践と普及に努めている<sup>2</sup>。

CPEは平成21年度に初めて群馬県安中市で導入され、東京都狛江市や岩手県においてもCPEの一部を適用した総合計画策定や政策評価が試行的に実施されている。

そこで本研究では、実際の行政に適用された CPEの実施前後において職員へのアンケート調査 を行い、その調査結果を分析することによって、 その有効性を検証する。さらに評価過程における 職場議論が活発であるかどうかがどのような因子 によって決定されるかを計量的手法によって明ら かにする。

## 2. CPEの概要

### (1) CPEの特徴

CPE は「自治体職員(又は市民等)が自らの 創造性(creativity)を発揮して、地域的公共的な 課題を解決するための政策の開発と創造を行う実 践志向型の政策評価である」と定義される(佐藤 2008)。このCPEには3つのタイプの評価(Ⅰ.現 場主導型、Ⅱ.政策選択型、Ⅲ.市民協働型)が 存在する。これらのうち本稿ではタイプ I を考察 対象とする。タイプ I は評価プロセスにおいて職 場討議を導入し自治体職員が「やりがい」「達成 感」を実感できるような対話型評価である³。

翻って、自治体で最も多く採用されているのが 事務事業評価である。尤も一口に「事務事業評価」 と言っても多彩多様であるが、CPEとの違いは概 ね次のように整理される(表1)。第1は評価目的 である。事務事業評価ではコスト削減、職員の意 識改革、説明責任の向上、効率的な行財政運営、市民参加の促進など多数の目的が並列されることが多い。一方、CPEの目的は「問題の解決」と「学習する組織への転換」である。前者は地域の公共的諸問題を発見し、それらの解決手段である事業の実施によって、どの程度解決されたか否かを検証することである。後者は組織の問題解決能力と、組織構成員である個々の職員の能力を開発し、組織の風土改革や生産性向上を目指すことである。

第2は意思疎通である。事務事業評価では、概して評価プロセスにおける職場討議の方法論が確立されておらず、評価シートの作成が目的化しがちである。CPEでは職場内に政策議論の場とプロセスを意図的に創出することによって小集団の討議による評価を実践する(ワークショップ方式の導入)。第3は主導部局である。事務事業評価では企画、財政、行革等の官房系部局が評価担当部局として主導的役割を担うが(集権型評価)、CPEではあくまで事業担当部局が主役であり、評価担当部局は事業担当部局の評価活動を支援する機能を担うにすぎない(分権型評価)。

第4は評価対象である。事務事業評価では通常数百本から数千本の事業を評価対象とする"広く浅い評価"であるのに対し(網羅型評価)、CPEでは政策的重要度の高い事業等に特化したメリハリのある評価を行う(選択型評価)。第5は評価頻度である。事務事業評価では同一の事務事業を毎年度評価するためマンネリ化しやすいが、CPEでは評価実施後は評価を踏まえて立案した事業計画を着実に実行することに力点が置かれる。この

表1	CPEと事務事業評価の対比
20	

	CPE	事務事業評価
主な目的	問題解決(政策の質的向上)	コスト削減(政策の量的削減)
土な日的	学習する組織への転換	説明責任の向上
意思疎通	職場討議	文書中心
主導部局	事業担当部局	官房系部局
評価体制	分権型	集権型
評価対象	重点的・深い	網羅的・浅い
評価頻度	必要に応じて適宜評価	同じ事業を毎年評価
指標データ	積極的に活用	参考程度

(注) ここでのCPEはタイプ I (現場主導の創造型政策評価) をさす。 (出所) 佐藤 (2008) を一部修正 ため評価自体は必ずしも毎年行わず、適宜評価を 行う(但し達成状況等を把握するためモニタリン グは毎年度行う)。

第6は指標データである。事務事業評価では指標やデータを参考程度に利用する場合も少なくないが、CPEでは事業への資源投入から成果発現に至る仮説を立て(ロジックモデルの作成)、問題(目標と現状とのギャップ)がどの程度解決したかを確認する際に積極的に活用する(仮説検証型評価)。

### (2) 評価ワークショップ

先述のように、CPEはいわゆる事務事業評価とは異なる特徴を有するが、最大の違いは職場における「評価ワークショップ」(以下「評価WS」とする)の存在である。評価WSでは「ファシリテーター」「事業説明者」「質問者」の3者が互いの役割を果たしながら、地域の公共的諸問題の解決に向けて職場討議による評価を行う。ここでは特にファシリテーターの役割が重要である。CPEのファシリテーターには課長級ないし課長補佐級職員1名があてられる。

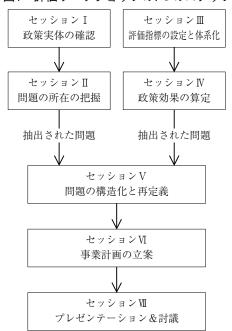


図1 評価ワークショップの7つのステップ

(出所) 佐藤 (2009)

ファシリテーターは、自身の部下や後輩を育成するという姿勢で、質問力、傾聴力、プレゼンテーション能力、論理的思考能力、評価指標設定技術、データ分析能力、問題発見能力、企画立案能力などの向上に努める。また自身も時間管理能力、リーダーシップ、質問力を磨くようにする。一般的にファシリテーターとはワークショップ参加者の意見を引き出し、議論を円滑に進める立場にある者をさすことが多いが、ここでは単なる司会者や進行役ではなく、CPEの推進リーダーとしてWS参加者の学習や気づきを促進する重要な役割を担っている。

この評価WSは図1に示した通り、7つのセッションから構成される。具体的には、①政策実体の確認、②問題の所在の把握、③評価指標の設定と体系化、④政策効果の算定、⑤問題の構造化と再定義、⑥事業計画の立案、⑦プレゼンテーションと討議、というよう問題の発見から解決に至る一連のプロセスを職場内で議論しながら段階的に評価を行う。特に(a)評価事業のロジックモデルを作成した上で、それを用いて直接成果、中間成果、最終成果が何であるかを徹底的に議論する、(b) それら成果を測定するのに適切な評価指標を議論によって導出する、(c) こうした過程を職場単位で実践する、という点に大きな特徴がある。

## 3. 先行研究の概観と本研究の特徴

自治体評価(行政評価)に関する先行研究は多岐に亘るため、本研究と同様に計量的な分析手法を用いた実証研究を中心に概観する。

湯下・坂野(2002)は行政評価システム導入先進県である三重県職員へのアンケート調査結果をもとにロジスティック回帰分析により、コスト意識、目的志向、顧客志向と関連する業務執行態度に対して各意識要因の単独効果を明らかにしている。古川・森川(2006)は、全都道府県を対象としたインベント・ヒストリー分析等から、評価が自治体から他の自治体へと波及し生成・実施されていく過程を明らかにしている。横山(2006)は、全国市区のクロスセッション・データをもとに被説明変数に行政サービス度を、説明変数に評価シ

80 佐藤 徹

ステムの導入・試行に関するダミー変数を投入し回帰分析を行い、行政サービスの向上に寄与する因子を析出している。また松尾(2009)は行政評価情報の利用に影響を及ぼしていると考えられる諸要因を研究仮説として導出し、これらを全国の都道府県及び市・東京特別区を対象に行った実態調査データを利用して検証している。

翻って、自治体評価の実務に目を転じれば、わが国の自治体評価は職員による自己評価をベースに行われている。事業担当職員による一次評価を起点として、さらに庁内での二次評価、三次評価を実施するなどしている。このうち本研究では一次評価プロセスに着目しており、既往研究の分析対象としてはそれほど多くない。

そもそも評価システムの導入当初は白紙の段階 から評価シートを作り上げなければならなかっ た。このとき評価シートの作成を通じて評価者が 個々の評価項目の記載内容を検討し、場合によっ ては職場での議論が期待されていたものと推察さ れる。だが導入後数年が経過すると評価シートに 記載すべき新規又は更新事項はかなり限定的とな り、職場組織での議論による検討に迫られること も少なくなる。また概して評価担当課は事業担当 課に対し、職場でよく議論して評価を行うように 依頼することはあっても、基本的にどのような手 順・方法で議論すればよいかについてはほとんど 示さず各職場の自主性に委ねることが多い。それ ゆえ評価過程における職場議論はあまり活発では ないう。本来は多様な活動であるべき評価活動が、 評価シートの作成に矮小化されている嫌いがある ことは否定できない(田中 2009)。その結果、職 場議論の成果として評価シートがあるのではな く、"よく見せるための文書づくり"に陥ってい たり、一部特定の職員が評価シートの作成を行い、 他の行政文書と同様に稟議による決裁行為のみで 意思決定してしまうことも少なくない。それでは たとえ同じ事業の担当者であっても評価活動に参 加しない職員は成果指標や目標数値に関心を持た ず、職場組織でそれらが十分に共有化されないこ とにもなりかねない。

かつて大森(1994)は、職場組織における政策 形成の潜在力を発揮していくためには、自由な発 議を奨励し政策形成課題の共有化を促進するため の「職場討議」を制度化することが不可欠である とした。そこで本研究では自治体職員の人材育成 や組織開発の観点から「評価」を捉え、これまで の評価研究ではあまり取り上げられてこなかっ た、評価過程における職場組織での討議に焦点を 当てることとした。

また研究の方法論上の特徴としては、第1に研究と実践の架橋を目指すアクションリサーチによって、自治体の評価担当課職員と協働してCPEの具体化とその実現に向けた制度設計を行ったり、筆者が導入教育のプログラムを作成し指導講師として参画したりするなど実践知の形式知への変換を試みようとしている点が挙げられる。第2はCPEを実際の自治体行政過程に適用し、職員へのアンケート調査によってその効果を検証しようとしている点である。第3は実務の中で行われる職員アンケートは一時点(事後)のデータによる推定評価が大半であるが、本研究では職員の意識や態度の変容を定量的に把握するため事前・事後の調査によりCPEの効果を統計的に分析しようとしている点が挙げられる。

## 4. 研究方法

### (1) 本研究のフィールド

本研究のフィールドは群馬県安中市である。同市を研究対象として取り上げたのは、①日本で初めて本格的にCPEを組織的に導入したこと、②筆者が同市の行政改革審議会行政評価部会長として行政評価システムの制度設計に参画する機会を得たこと、③市の職員(実務家)と協働して改革を行い、その過程で理論の実践と検証を行える協力体制が得られたこと、などの理由による。

安中市は群馬県南西部にある人口約6万1千人、 面積約276平方キロの地方都市である。平成18年3 月、旧安中市と旧松井田町が新設合併し、現在の 人口と市域になった。同市では平成18年6月に、 地方自治法第138条の4第3項の規定に基づく行政 改革審議会を設置した。同審議会は、合併後の行 政改革大綱の策定など行財政の幅広い課題につい て調査・審議・答申を行う附属機関である。委員 は学識経験を有する者(第1号委員)、公共的団体 を代表する者(第2号委員)、その他市長が必要と 認める者(第3号委員)の各5名、計15名で構成さ れている。

平成18年12月4日、同市は行政評価のあり方を 含む6項目(①安中市行政改革大綱(集中改革プ ランを含む)の策定、②行政評価のあり方、③補 助金のあり方、④市有財産の有効活用、⑤公の施 設に係る指定管理者制度の運用、⑥使用料、手数 料のあり方)を行政改革審議会に諮問した。これ らのうち当初審議会の検討事項は行政改革大綱 (集中改革プランを含む) の策定に主眼が置かれ た。およそ1年半かけて審議がなされ、平成20年3 月に策定された『安中市行政改革大綱(平成20年 度~平成26年度)』において、「限られた資源の中 で市民サービスの質的向上や行政事務の効率化・ 高度化を図るため、行政評価を効果的・積極的に 活用するとともに、行政運営へのチェック機能の 充実を進めていくこと」が明記された。こうして 同市において行政評価制度の導入に向けた検討が 進められることになった。そして具体的な検討は、 行政改革審議会のもとに設置された行政評価部会 において行われた6。同部会では「そもそも行政 評価は手段であって目的ではない。それゆえ行政 評価の制度設計ではその目的設定が極めて重要と なる。目的が多数羅列されると、結局何を目指す かが不明瞭となり、目的自体を見失いがちである | とし、「先行自治体の多くが陥った轍を踏まない」 「安中市ならではの行政評価を目指すこと」が部 会で確認された。特に行政評価が事業担当部局に とっての"余計な仕事"と捉えられることがない よう、職員の能力向上と意識改革に重点を置いた 評価制度を設計するものとされ、CPEの導入が了 承された。

以上をふまえ、平成21年2月4日、行政改革審議会は岡田義弘市長に対して、「行政評価は、今後市民サービスの提供を進めていく上で、市民にとっての成果に着目した真に効果のあるものとすること、また市政の運営を進めていく上で、組織及び行政運営の改革、人材の育成を図り、効率的なものとすることを基本として取り組むことが必要である」と答申した。これを受けて、平成21年5月、岡田市長は「行政評価の導入に伴う研修の実施について」と題する訓示文書の中で、平成21年

度から行政評価の導入を行うことを宣言した。

### (2) CPEの導入実践

### ①導入教育の実施

「今なぜCPE なのか」「他の自治体の事務事業 評価と比べてCPEはどこが違うのか」「CPE によってどのような組織文化への変革を目指そうとしているのか」。こうしたCPE のコンセプトを組織全体で共有化することは導入初期には不可欠である。

そこで、全職員がCPEの意義等を理解し一体となって推進することができるよう、市長から部課長を通じ各部署の職員に研修への参加が命じられた(表2)。

## ②評価事業の選定

CPE では、いわゆる事務事業評価のように評価対象を網羅的に広く設定するのではなく、範囲を絞り込み評価を行う。また導入教育受講者がファシリテーターとなって、自身の課が所管している事業群の中から、当該年度にどの事業の評価に取り組むべきかを課内論議する。ただし安中市ではCPE導入初年度であることに配慮し、職員への意識啓発も兼ねて1係1事業を基本としてCPEの評価事業を100事業に絞り込むことにした。

またCPEでは政策的に重要な事業(総合計画の 重点施策に位置付けられた事業、実施計画の主要 事業等)、長期間実施してきた継続事業で何らか の問題が生じていそうな事業、課や係の基幹事業、 事業費の大きな事業などに絞り込み評価を行う が、この点についても導入初年度である点に鑑み、 各担当課が「改善・改革の余地が高く、所管課・ 係としても改善・改革したいと考えている事業」 を1つ以上最大3つまで選定し、その情報をもとに 企画課と協議しCPEの評価事業を決定した。

### ③評価チームの編成

職場討議を行う評価チームは評価事業の担当係を中心にして1チーム5名程度で編成するものとした。

ファシリテーターは養成研修を受講した者があたり、また事業説明者には評価事業の内容を最も 熟知している担当者1名(担当係長又は主査)が 82 佐藤 徹

<b>=</b> 0	ODEを中吐ナッ	ための導入教育
オマン	じアトを手託する	) 7.7 (8) (7) 2基 人 28 百

		導入研修	実務研修	ファシリテーター 養成研修
開催日君	д	平成21年6月2日、3日、	平成21年7月7日、8日	平成21年8月11日、12日、
用惟口怕	生.	11日、12日		17日、18日
対 象 者	by.	市長、部長をはじめとする各	課長職・係長職に加えた各係	係長職から選抜された職員
<b>刈</b>	3	階層職員(計416名)	1名程度(計220名)	(計64名)
		職階別に経営者研修(市長、	職場議論を活性化する際に欠	1回の研修は連続2日間として
		部長級職員)、管理職研修(課	かせないファシリテーターの	計4日間実施された。研修の
		長級職員)、一般職研修(係	役割や評価ワークショップな	内容は、講義と演習(評価ワ
		長級以下職員)の3つで構成。	ど CPE の全体像が提示され	ークショップ形式)で、筆者
	容	導入研修では主に行政経営の	た。また直接成果指標、中間	が講師は務めた。1回(2日間)
		基本概念、行政評価の導入背	成果指標、最終成果指標など	の研修では全体の講義後(0.5
内 容		景とその必要性など基礎的知	の各種指標の設定方法等につ	日)に、5チーム(1チーム当
n a		識の修得を目的として行われ	いての理解を深めるため、	たり5~6名)に分かれ、1チ
		た。他団体で導入されている	「ハザードマップ作成配布事	ームにつき市で実際に行われ
		事務事業評価とは違い、CPE	業」と「駅前自転車駐輪場整	ている事業を1つずつ取り上
		は職場議論の活性化、職員の	備事業」の2事業を取り上げ、	げ、評価ワークショップが行
		モチベーションや政策能力の	簡易なワークショップが実施	われた(1.5日)。ファシリテ
		向上に力点があることが説明	された。	ーターは企画課が予め指名し
		された。		た。

(注) 評価WSで取り上げた計10事業は、(イ) 事業がソフト系、ハード系がどちらか一方に著しく偏らないこと、(ロ) 事業名からある程度内容がイメージできること、(ハ) 内部管理ではなく、市民を対象とした事業であること、(ニ) 事業が特定の部署に偏らないようにすることを考慮し、筆者が選定した。

(出所) 筆者作成

あてられた。その他のメンバーはすべて質問者として参画するものとした。その際できるかぎり当該事業を直接担当していない職員(別の課の職員でもよい)を数名加えるものとした。担当者とは違った視点での評価が期待でき、組織内の情報共有も図られるからである。

### ④議論の場の設定とフォローアップ

ファシリテーター養成研修では実際の10事業を評価対象としたが、残りの90事業については受講者が研修終了後に各職場においてファシリテーターとなり、評価WSを実践することが企図されていた。しかし各職場で評価WSが即座に実行できるかといえば、そう簡単ではない。すべての受講者が一様にファシリテーターとしてのスキルを習得しているとは限らず、また研修の受講対象とならなかった職員も存在していることが想定されたからである。

そこで、CPEの7つのセッションのうち、セッションⅢ(評価指標の設定と体系化)を中心に、ロジックモデルの作成方法や指標設定のスキルの

習得を目的とした集合型評価ワークショップ研修が実施された。同研修は残りの90事業のうち30事業を題材として、当該事業を担当する部署の係長や係員を対象に行われた\*。その後、企画課は事業担当課に対して、議論の場を設定してもらうよう積極的に働きかけるとともに、ファシリテーターの補佐役としてワークショップにできる限り参加したり、ヘルプデスクを設けた。このようにして30事業に関してはセッションⅣ以降を、また残り60事業についても同様に各職場で評価WSが行われた°。

### (3) 調査方法

安中市の全職員480名(病院医事課を除く)に対して、無記名の自記式質問紙による調査を行った<sup>10</sup>。CPE(職場討議による評価)を実施する前(平成21年10月)と実施した後(同年12月)の2回にわたって同一項目による調査を行った。この事前・事後の調査項目はCPEの実施によって、(A)職場活性化状況、(B)行政評価への関心や必要性、(C)評価技術の習得度、(D)業務に取り組

## 表3 CPE実施前後の共通質問項目

Q1	職場 (課レベル) では、組織の目的やビジョンが明確 に示されている
Q2	職場(課レベル)では、組織の目的やビジョンの実現に向け、メンバー(所属職員)が力を合わせて貢献している
Q3	職場 (課レベル) では、情報が共有化され、円滑なコミュニケーションが図られている
Q4	事業の目的や成果が職場 (係レベル) で共有化されている
Q5	事業の問題点や改善策が職場(係レベル)で共有化されている
Q6	行政評価に関心がある
Q7	行政評価の意義や必要性を認識している
Q8	住民の目線に立って、事業の目的や成果を考えている
Q9	成果に見合った事業費 (コスト) となっているかどう かを考えている
Q10	行政評価の目的を理解している
Q11	行政評価に必要な論理的思考方法を習得している
Q12	行政評価における職場議論の方法を習得している
Q13	評価指標の導き方を習得している
Q14	行政経営(人、金、物の効率的活用)を担っているという意識で仕事にあたっている
(注)	- 久頂日の日度けすべて「士いにそう田う」と「今くそう田わた

(注) 各項目の尺度はすべて「大いにそう思う」と「全くそう思わない」を両端とする5段階評価とした。

(出所) 筆者作成

む姿勢、などに変化が見られるのではないかという仮説のもと、表3に示す14項目を設定した"。各項目は全て「大いにそう思う」「ややそう思う」「どちらでもない」「あまりそう思わない」「全くそう思わない」の5件法による質問方式とした。

さらに、事後アンケートでは「職場での討議に よる評価」の実施状況やそれ自体の評価を尋ねた。 ここでの調査項目は表4に示す10項目である。同 様に5件法による質問手法を採用した。

なお、属性及び基礎情報として職階、所属名、 年齢、公務員としての職務経験年数、入庁時の団 体、研修の受講暦、評価ワークショップでの役割 を取り上げた。

### (4) 分析方法

## ①CPEの効果分析

CPEの効果を把握するために、職員に対するアンケート調査によって、(A)職場活性化状況、(B)行政評価への関心や必要性、(C)評価技術の習得度、(D)業務に取り組む姿勢などの諸点

### 表4 評価WS参加者への質問項目

QA1	職場での評価はおおむね手引き(「創造型政策評価 (CPE) ワークショップの進め方』や『安中市行政評価: 創造型政策評価 (CPE) 職場マニュアル』)の内容に沿って行われた
QA2	職場では活発に議論しながら評価が行われた
QA3	ファシリテーターは議論の場を適切にマネジメントした
QA4	行政評価に対する上司の関心は高いほうである
QA5	企画課や支援業者 (コンサル) のサポート体制が整っ ていた
QA6	事業の見直しに有意義であった
QA7	今後も継続実践することは、組織風土の改善改革に役立つ
QA8	今後も継続実践することは、職員の政策能力の向上につながる
QA9	事業に関する様々な問題と情報を,職場で共有できる ことを実感した
QA10	評価ではシート (調書) づくり以上に, 議論の場とプロセスが重要であることがわかった

(注) 各項目の尺度はすべて「大いにそう思う」と「全くそう思わない」を両端とする5段階評価とした。

(出所) 筆者作成

について、CPEの実施前後で測定し、その回答に 差があるかどうかについて統計的に検定(t検定) を行った。

## ②評価プロセスにおける職場議論活性化の規定要 因の分析

評価プロセスにおいて職場議論が活発に行われるかいなかはどのような因子によって決定されるのだろうか。これを明らかにするため、AIC統計量(Akaike's Information Criterion;赤池情報量規準)によって計量分析を行った。AIC統計量は次式により計算される $^{12}$ 。

AIC= (-2) Log (最大尤度) +2 (自由パラメータ数)

## 5. 分析結果と考察

### (1) 分析対象者の基本属性

事前調査では435名から回答があり(回答率90.6%)、事後調査では439名から回答があった(回答率91.5%)。これらのうち事前・事後の質問紙の両方に回答し欠損値のない者174名を有効回答とし、分析を行った。

分析対象者の属性を表5に示す。回答者の職階

表5	分析対象者の基本属	性
1×J		1 —

n=174 上段% 下段人数

							工权	%,下权八奴
	課長	係長	主査	主任	主事			
職階	2.9	39.7	30.5	9.8	17.2			
	5	69	53	17	30			
	25歳未満	25~29歳	30~34歳	35~39歳	40~44歳	45~49歳	50~54歳	55歳以上
年 齢	3.4	5.7	13.2	12.6	10.3	20.7	21.3	12.6
	6	10	23	22	18	36	37	22
経験年	5年未満	5~9年	10~14年	15~19年	20~24年	25~29年	30~34年	35年以上
数数	8	10.3	9.8	13.2	16.7	21.8	11.5	8.6
**	14	18	17	23	29	38	20	15
入庁時	旧安中市	旧松井田町	新安中市	その他				
の団体	53.4	28.2	5.2	13.2				
♥/国件	93	49	9	23				
WSでの	事業説明者	ファシリテーター	質問者					
役割	17.8	25.3	56.9					
12.01	31	44	99					

(注) 課長には参事が、係長には主幹・課長補佐が、主査には係長代理が、主事には技師・主事補・技師補がそれぞれ含まれる。 (出所) 筆者作成

は、係長(主幹、課長補佐を含む)が39.7%と最も多く、次いで主査(係長代理を含む)30.5%であった。その他、主事(技師、主事補、技師補を含む)17.2%、主任9.8%、課長(参事を含む)2.9%であった。

年齢構成を見ると、 $50\sim54$ 歳が最も多く約2割を占め、これに $45\sim49$ 歳を含めると4割以上を占めた。20代は9.1%と1割に満たなかった。また公務員としての経験年数は、 $25\sim29$ 年が最も多く、全体の21.8%であった。

入庁時の団体は、合併前の旧安中市が53.4%、旧松井田町が28.2%であり、合併後の新安中市に採用された者は5.2%であった。その他(事務組合等)の採用は13.2%であった。評価ワークショップでの役割としては、事業説明者17.8%、ファシリテーター25.3%、質問者56.9%であった。

## (2) CPEの効果

表6にCPEの実施前後における14の質問項目の 平均スコアと標準偏差、及び事前事後のt検定 (両側検定)の結果を示す。

この分析結果からは、「行政評価の目的を理解している」(Q10)、「行政評価に必要な論理的思考方法を習得している」(Q11)、「行政評価における職場議論の方法を習得している」(Q12)、

「評価指標の導き方を習得している」(Q13)の4項目について平均スコアの上昇がみられ、CPEの実施前後で統計的に有意な差が確認された(p<0.05)。このうち、特に「論理的思考方法」(Q11)、「職場議論手法」(Q12)、「評価指標導出」(Q13)の3項目についてはいずれもCPEが最も中心的な教育効果として掲げているものである。

そして、「行政評価に関心がある」(Q6)、「行政評価の意義や必要性を理解している」(Q7)、「住民の目線に立って、事業の目的や成果を考えている」(Q8)、「成果に見合った事業費(コスト)となっているかどうかを考えている」(Q9)、「行政経営(人、金、物の効率的活用)を担っているという意識で仕事にあたっている」(Q14)といった行政評価又は行政経営に関する項目に関しては、統計的に有意な差は見られなかった。

さらに組織や職場の意思疎通に関する項目 (Q1、Q2、Q3) についてもCPEの実施前後で有意な差を確認することができなかった。これらはもともと高い平均値を示しており、これ自体が組織風土や組織文化に係わる問題であることに鑑みれば、一度のCPEの実施で大きく改善されるというよりもむしろ今後CPEを継続実施していく中でそうした効果が得られるのではないかと考えられる。このほか、「事業の目的や成果が職場(係レベ

表6 効果の検定

n=174

		Ī	事前	Ē	事後		t検定	
		平均	標準偏差	平均	標準偏差	t値	自由度	p値
Q1.	組織ビジョン明示	3.70	0.98	3.78	0.92	-1.15	173	0.251
Q2.	組織メンバー協力	3.79	1.00	3.82	0.89	-0.34	173	0.731
Q3.	コミュニケーション	3.71	0.97	3.59	0.96	1.75	173	0.082
Q4.	事業目的共有化	3.84	0.92	3.86	0.86	-0.28	173	0.777
Q5.	事業問題共有化	3.76	0.93	3.74	0.84	0.29	173	0.775
Q6.	行政評価関心	3.35	0.96	3.34	0.92	0.17	173	0.862
Q7.	行政評価必要性	3.60	0.93	3.66	0.87	-0.98	173	0.330
Q8.	住民目線立脚	4.02	0.68	4.03	0.68	-0.24	173	0.814
Q9.	費用対効果	3.70	0.80	3.77	0.80	-1.20	173	0.230
Q10.	行政評価目的理解	3.60	0.82	3.82	0.70	-3.59	173	0.000 **
Q11.	論理的思考方法	2.66	0.84	2.93	0.83	-5.17	173	0.000 **
Q12.	職場議論手法	2.55	0.81	2.90	0.81	-5.81	173	0.000 **
Q13.	評価指標導出	2.46	0.86	2.77	0.85	-5.35	173	0.000 **
Q14.	行政経営意識	3.80	0.87	3.75	0.79	0.73	173	0.467

\*\* p < 0.05

(出所) 筆者作成

ル)で共有化されている」(Q4)と「事業の問題点や改善策が職場(係レベル)で共有化されている」(Q5)についても有意な差が見られなかった。これらについてはなぜ有意差が見られなかったかは不明であるが、職場内で職員が所管事業の中から評価対象事業を選定する過程で議論を行うため、Q4やQ5といった効果が得られた可能性が考えられる。

## (3) 評価プロセスにおける職場議論活性化の規定要因

つぎに、評価過程において職場議論が活発であるかどうかがどのような因子によって決定されるか、その規定要因の分析結果を示すことにする。

表7は、職場議論活発度(QA2)を目的変数とし、職員の属性(職階、年齢、WSでの役割)及び前述のQ1~Q14及びQA1~QA10の各項目を説明変数としたときの、説明変数の有効性の大きい順に上位20位までリスト化したものである。ここではAICの値が小さいものから順に整列させている。第1位から第9位まではAIC値がマイナスであるため、説明変数と目的変数の関係は独立ではない、すなわち統計的に有意であると解釈できる。これらのうち以下では上位3位までみてみよう。

まず説明力の最も大きな変数(第1位)は、「フ

ァシリテーターによる議論の場のマネジメントの適切度」(QA3)であった。これはAICが-62.78と最も小さい。つまり、職場で活発に議論しながら評価が行われたかどうかは、ファシリテーターが議論の場を適切にマネジメントしたかどうかによって最も大きな影響を受けると解釈できる。

表8は「職場議論活発度」(QA2)と「ファシリテーターのマネジメントの適切度」(QA3)との間のクロス表であるが、これからも明らかなように、「ファシリテーターのマネジメントの適切度が高い」と評価した職員はその96.6%が「職場では活発に議論しながら評価が行われた」としている。他方、「ファシリテーターが議論の場を適切にマネジメントした」とは思わなかった職員はその86.6%が「職場では活発に議論しながら評価が行われた」と思っていないということがわかる。

そして説明力が2番目に大きかった変数は「職場議論評価ルール」(QA1)であった。つまり、CPEの進め方をまとめた手引き書(マニュアル)に沿って職場での評価WSが行われたかどうかが、評価過程における職場議論の活発度にかなりの影響を与えていることがうかがえる。

さらに説明力の大きい項目の第3位は「行政評価上司関心」(QA4)であった。評価に際して職場での議論が活発に行われるかどうかは、職場の

86 佐藤 徹

表7 評価プロセスにおける職場議論の活性化の規定要因

順位	説明変数	AIC統計量	AIC値の差	v <sup>2</sup> 統計值	P値
//只元 ————	10年7月交换	7110/元日至	THCEOE	χ <sup>ηγι</sup> ΠΠΕ	1 112
1	ファシリテーター(QA3)	-62.78		105.33	0.00
2	職場議論評価ルール (QA1)	-48.50	14.29	89.64	0.00
3	行政評価上司関心(QA4)	-37.61	10.89	76.68	0.00
4	事業目的共有化(Q4)	-19.14	18.47	52.30	0.00
5	事業問題共有化(Q5)	-14.69	4.45	51.97	0.00
6	事業見直し意義(QA6)	-8.28	6.41	35.19	0.00
7	住民目線立脚(Q8)	-7.50	0.79	31.63	0.00
8	職場共有化実感(QA9)	-6.68	0.82	34.37	0.00
9	職階 (F1)	-1.72	4.96	26.41	0.01
10	組織ビジョン明示(Q1)	0.93	2.64	25.21	0.01
11	行政評価関心(Q6)	1.76	0.84	24.04	0.02
12	原課サポート体制(QA5)	3.13	1.36	26.33	0.01
13	評価指標導出(Q13)	5.13	2.00	13.25	0.15
14	組織風土改善改革(QA7)	6.26	1.13	18.13	0.11
15	組織メンバー協力 (Q2)	6.41	0.15	21.53	0.04
16	役割 (F3)	7.12	0.72	4.53	0.61
17	費用対効果(Q9)	8.48	1.35	11.50	0.24
18	議論の場とプロセス (QA10)	8.63	0.16	19.98	0.07
19	行政経営意識(Q14)	8.89	0.26	9.97	0.35
20	コミュニケーション (Q3)	10.33	1.44	16.04	0.19

(出所) 筆者作成

表8 評価プロセスにおける職場議論の活性化とファシリテーターの役割

職場では活発に議論しながら評価が行われた ファシリテーターは 議論の場を適切に マネジメントした	大いにそう思う	ややそう思う	どちらでもない	あまりそう思わない	全くそう思わない	合計
大いにそう思う	16 53.3	13 43.3	3.3	0.0	0.0	30 100.0
	10	54	9	2	0.0	75
ややそう思う	13.3	0.7	0.1	0.0	0.0	1.0
104 2 - 1 4	3	16	24	5	0	48
どちらでもない	6.3	33.3	50.0	10.4	0.0	100.0
ままりてき田ねむい	0	3	4	8	0	15
あまりそう思わない	0.0	20.0	26.7	53.3	0.0	100.0
全くそう思わない	0	1	3	2	0	6
主くてリ芯わない	0.0	16.7	50.0	33.3	0.0	100.0
△ ⇒+	29	87	41	17	0	174
合 計	16.7	50.0	23.6	9.8	0.0	100.0

(注) 上段が人数、下段が%

(出所) 筆者作成

上司が行政評価にどの程度関心を持っているかに も大きく依存することが示唆された。

## 6. 結論と今後の課題

行政学では、かねてより自治体職員の人材育成や組織開発の視点から職場組織での討議の必要性が指摘されていた(大森 1994)。そこで本研究では職員の人材育成や組織活性化の観点から行政評価を捉え、既往の評価研究ではあまり取り上げられることがなかった、評価過程における職場討議に着目し、その方法論を体系化したCPEを現実の行政過程に導入し効果の検証を試みた。

その結果、CPEを実践することによって、職員は評価の目的を理解したり、評価に必要な論理的思考方法や評価指標の導出方法が習得できるようになり、職場で議論しながら評価できるようになる可能性が示唆された。このような評価技術の習得度に関する効果については確認されたが、職場組織の活性化、評価への関心度、業務に取り組む姿勢の変化などの効果については確認できなかった。これらはどちらかというとCPEの単年度の実施によって得られる効果というよりは一定期間長期にわたりCPEを継続実施していくことによって得られる効果であると考えられる。

また、本研究では評価過程における職場議論が活発であるかどうかがどのような因子によって決定されるかについて分析した。その結果、職場で活発に議論しながら評価が行われるかどうかは、ファシリテーターが議論の場を適切にマネジメントしたかどうかに最も大きく依存していることが明らかとなった。この点については、単に行政評価の理念や手法の習得を目的とした講義型研修ではなく、ファシリテーターの養成を目的としたグループワークによる演習形式の職員研修を恒常的に行う必要があるだろう。

なお、本研究ではCPEの有効性を職員に対するアンケート調査によって把握しようとした。CPEの実施による職員の意識変化や学習効果がこの評価制度や評価手法の有効性を高める可能性があると考えたからである。しかし、アンケート調査には評価指標やロジックモデル等に関する知識を問

う設問が含まれなかったため、回答者の主観的評価となっている。この点は職場議論の活性化に関するアンケートについても同様であり、今後はより客観性を高めるための研究上の工夫が求められる。また、本研究ではCPEを我が国で初めて本格的に導入した安中市の職員アンケート調査に基づく実証研究であったが、今後ともCPEを継続的に実施することによって、それが職員の心理や行動、さらには組織風土にどのような影響を及ぼすのかについても経過観察していくことが残された今後の課題である。

#### 謝辞

本研究の遂行にあたり、調査に御協力を頂いた 安中市の企画課をはじめとする職員の皆様に深く 感謝の意を表します。また匿名の査読者の方々か らは、大変有意義なコメントを頂きました。心か ら感謝申し上げます。

### 注記

- 1 山谷(2008)は自治体評価の10年を振り返り、4つの類型を示している。それらのうち、評価が重要なツールであると認識しているが担当者の基本的な態度としては「導入すればよい」と考えていた自治体を第3の類型としており、こうした自治体では評価が形骸化しマンネリ感に沈むことになると指摘している。
- 2 CPEのタイプ I は高崎経済大学附属地域政策研究センターにおいて、自治体職員を対象として平成13 年度より6年間毎年開催された「自治体政策評価演習」の実践的経験と成果によるところが大きい(北海道から沖縄まで全国各地から約178団体、計244名の自治体職員が参加)。同演習は斎藤達三氏(高崎経済大学名誉教授)と筆者を指導講師として実施されたが、平成19年度から「職場議論活性化のための創造型政策評価・実践演習」としてリニューアルされ、平成21年度まで筆者が指導講師をつとめた。
- 3 Hackman & Oldham (1976) は仕事の内容的特性が 動機付けに大きな影響を与えるという職務特性モデ ルを提唱している。評価自体を職務と捉えた場合、 CPEにはHackman & Oldham (1976) がいう5つの職

88 佐藤 徹

務特性(技能多様性、仕事の完結性、仕事の有意義性、自律性、フィードバック)が具備されており、評価に対する動機付けや満足感だけでなく、生産性の向上が期待できる(佐藤 2008)。

- 4 わが国においてワークショップは1990年代半ばから 都市計画における住民参加の場で適用され、その後 計画や条例づくりなど幅広く行われている (伊藤・原田 2011)。CPEではワークショップを職員参加の 評価過程に適用したものである。評価ワークショップの詳細については佐藤 (2008) の第4章「政策評価ワークショップ」を参照されたい。
- 5 筆者が全国の市・特別区 (806団体)を対象に実施した『自治体における行政評価と行政経営に関する実態調査』(平成21年3月-4月調査、回収率69.9%)によれば、「全庁的に職場で活発に議論しながら行政評価が行われていますか」との設問に対して「はい」と回答した団体はわずか14.4%に過ぎなかった。
- 6 行政評価部会のメンバーは筆者以外に、中小企業診 断士、公認会計士、群馬県西部県民局長、元安中市 財政課長で構成されている。
- 7 CPE導入初年度である平成21年度には予算事業、合併時のすり合わせ事業、実施計画の主要事業のデータ等を参考に、事務事業棚卸シートを使い1834事業が棚卸しされた。これをベースにCPEの対象事業が絞り込まれた。なお平成22年度についてはCPE評価対象事業をさらに絞込み60事業に選定している。
- 8 平成21年10月5日、7日、8日の3日間 (9:00 ~ 11: 30、13:15 ~ 15:45の計6コマ) 実施された。
- 9 その後CPE の評価情報は事業マネジメントシート として係長以下がとりまとめ、課長・部長決裁によ る1次評価が行われた。さらに庁内組織である行政 改革推進会議が2次評価を行い、予算編成に活用す るものとされた。
- 10 質問紙は筆者が作成したが、質問紙の配布回収は安中市企画課の協力を得た。また回答結果は統計的に 処理するため、個人情報は公表されないことを紙面 に明記し、倫理的配慮を行った。
- 11 (A)、(B)、(C)、(D) に関する質問としてそれぞれQ1・2・3・4・5、Q6・7、Q11・12・13、Q8・9・14が該当する。なお(A) に関する質問項目を検討するにあたっては、高橋(1989)の「組織が活性化された状態」の定義、すなわち「組織のメンバーが、①組織と共有している目的・価値を、②能動的に実現して

いこうとする状態 | を参考にした。

12 AIC統計量を用いる最大のメリットは、極めて広範なデータの中から、質的な目的変数に対して最も多くの情報を持つ説明変数を自動的に探索するためのプログラムであるCATDAP(CATegorical Data Analysis Program)(Katsura and Sakamoto1980)を利用できる点にある。CATDAPは目的変数と全説明変数との間のあらゆるクロス表を作ってAICの値を計算し、有効度の大きい順に説明変数をリスト化したり、説明変数の最適な組合せを検出することができる(小野 1995)。またこのプログラムは、目的変数が質的な変数であること以外に、適用上の制限条件はなく、説明変数の数、種類、サンプルサイズの如何によらず適用できることも、AICを用いた理由である(坂元 2010)。

## 参考文献

伊藤雅春・原田和成 (2011)「市民参加手法の比較検討」、『コミュニティ政策研究』、13:1-23

大森彌 (1994) 『自治体職員論』、良書普及会

小野能文(1995)「多変量解析法としてのCATDAPについて」、『夙川学院短期大学研究紀要』、19:63-79

斎藤達三 (2001) 『自治体政策評価演習』、ぎょうせい 坂元慶行 (2010) 「統計的日本人研究雑感-ある国民 性調査係の36年の思い出」、『統計数理』、58(1): 61-82 佐藤徹 (2008) 『創造型政策評価』、公人社

- 佐藤徹(2009)「創造型政策評価の実践法 第3回:職場 議論の中から問題を抽出しよう!」、『月刊地方自治 職員研修』、公職研、42(9):82-83
- 高橋伸夫(1989)『組織活性化の測定と実際』、日本生 産性本部
- 田中啓 (2009)『日本の自治体の行政評価』(分野別自 治制度及びその運用に関する説明資料No.14) 財団 法人自治体国際化協会、政策研究大学院大学比較地 方自治研究センター
- 田渕雪子(2010)「地方自治体における行政評価12年 の歩みと今後の展望」、『三菱総合研究所所報』、53: 30-53
- 藤井達也(2006)「参加型アクションリサーチーソーシャルワーク実践と知識創造のために」、『社会問題研究』、55(2):45-64
- 古川俊一・森川はるみ (2006) 「地方自治体における

評価の波及と生成過程の分析」、『日本評価研究』、6 (1):133-146

松尾貴巳 (2009) 『自治体の業績管理システム』、中央 経済社

山谷清志 (2008) 「特集「自治体評価の10年」について |、『日本評価研究』、8 (1):1-2

湯下健一・坂野達郎 (2002)「行政評価システム導入 県における目的指向型行財政運営の実態と評価」、 『計画行政』、25 (1):109-117 横山麻希子 (2006)「評価による行政サービスの向上」、 『日本評価研究』、6 (2):59-71

Hackman, J.R., & Oldham, G.R.(1976). Motivation through the design of work: test of a theory. *Organizational behavior and human performance*, 16 (2): 250-279.

McNiff, J.and Whitehead, J. (2006). *All You Need to Know about Action Research*, London:SAGE Publications.

(2011.8.17受理)

# A Empirical Study of the Effects of Discussions in the Workplace on the Evaluation Process

Toru Sato

Takasaki City University of Economics tsato@tcue.ac.jp

### Abstract

Policy evaluation was introduced in Japanese local governments starting in the mid-1990s. Although evaluation appears to be an established practice to certain extent now, staff in charge of evaluation continue to feel evaluation is a forced obligation and burdensome. To address these issues the author proposes Creative Policy Evaluation (CPE). In this investigation, CPE was introduced into administrative practice and its effects were verified through staff questionnaire. As a result of the introduction of CPE, a statistically significant difference in the comprehension of the purpose of evaluation, achievement of the logical thinking, derivation of the performance indicators, and discussions in the workplace in an evaluation process were seen. By analyzing the activeness of the discussion in an evaluation process by related factors it becomes clear that actiiveness in discussion is dependent on appropriate management of the discussion by the facilitator.

## Keywords

local government evaluation, workplace dialogue, evaluation workshop, Creative Policy Evaluation

## 【研究論文】

## 技術協力プロジェクトの効果に及ぼすガバナンスの影響

## 菊田 怜子

牟田 博光

東京工業大学大学院 社会理工学研究科 kikuta.r.aa@m.titech.ac.jp 東京工業大学大学院 社会理工学研究科 muta@hum.titech.ac.jp

## 要約

開発援助プロジェクトの効果に及ぼす被援助国のガバナンスの影響に関する多くの研究がされるとともに、ガバナンスの状況を測る様々な指標が考案されている。日本のODAプロジェクトの効果と被援助国のガバナンスとの関連を定量的に分析した研究はこれまでされてきてない。本研究はJICAが実施した181の技術協力プロジェクトを対象としてプロジェクト効果に対するガバナンスの影響について因果モデルを構築して分析した。ガバナンス指標にはWBIの世界ガバナンス指標を用いた。その結果、ガバナンスはプロジェクト効果の発現に対しては影響を及ぼさないが、プロジェクト終了後における効果の活用・持続性に対して影響を及ぼすことが明らかになった。さらに、プロジェクト効果に対するガバナンス、当該国の経済力、教育の影響を検証したところ、ガバナンスはプロジェクト効果の活用・持続性に直接影響を及ぼすのではなく、経済や教育を通して間接的に影響を及ぼすことが明らかになった。

## キーワード

ガバナンス指標、技術協力プロジェクト、援助効果、共分散構造分析

## 1. はじめに

開発途上国が自立し、持続的に発展するための 貧困削減にむけて、様々な試みによる援助が行わ れている。援助の効果が発現した国、あるいは発 現しなかった国があるが、援助の効果には援助国 による努力ばかりでなく受け手である被援助国側 の努力が影響を与えることは否めない。

被援助国において、経済が良く管理されており 且つ援助を多く受けている国は経済の成長がみられる(World Bank 1998)、経済政策が良い国では 乳児死亡率が低い(Burnside and Dollar 1998)、援助は成長に寄与するが、熱帯地域では寄与しない (Dalgaard, et al. 2004)、健全な経済政策をとる国は援助効果が現れる (Burnside and Dollar 2000、2004)、などの研究が示すように、良いガバナンス (グッド・ガバナンス) によって援助が効果的に現れることがいわれている。しかし、その一方、被援助国への援助が多いとガバナンスの質が悪くなる (Knack 2001, Knack and Rahman 2007)、あるいはガバナンスが弱体化する (Rajan and Subramanian 2007) という研究が示すように、被援助国のガバナンスと援助には関連性がみられ、開発援助におけるグッド・ガバナンスの重要性が増している。

ガバナンスの定義は様々で、その概念の解釈も 援助国によって異なるという多様性をもっている が (下村 1999、近藤 2003、国際協力機構 2008)、 主として"行政機能の効果・効率や透明性・説明 責任、政権・政策の手続き的な正統性、法の支配、 関係者や社会の構成員の意思決定や執行への参 加"という点から論じられることが多い(稲田 2006)。国際協力機構(2008)ではガバナンスを、 "ある国の安定・発展の実現にむけて、・その国 の資源を効率的に、また国民の意思を反映できる 形で、動員し、配分・管理するための政府の機構 制度、・政府・市民社会・民間部門の協働関係や 意思決定のあり方、など、制度全体の構築や運営 のあるべき姿"と定義している。また、2003年の 政府開発援助大綱では基本方針の一つとして良い 統治(グッド・ガバナンス)に基づく開発途上国 の自助努力を支援することをあげている(外務省 2003)。"その際、平和、民主化、人権保障のため の努力や経済社会の構造改革に向けた取組を積極 的に行っている開発途上国に対しては、これを重 点的に支援する"としており、日本の開発途上国 への援助にあたって被援助国のグッド・ガバナン スを重視している。

開発援助におけるガバナンスの重要性が増すとともに、被援助国のガバナンスを考慮した援助が行われるようになってきた。ではガバナンスの状況をどのように評価しているのであろうか。多くの国をカバーし、一般的に用いられているガバナンス指標には次のようなものがある。

## ・フリーダムハウス:民主化指標(DI: Democracy Index)

194ヵ国3地域(Freedom House 2009)を対象として、国内外の情報、NGO出版物、シンクタンク・学術分析等による情報を二次データとして用いており、専門家の主観的認識に基づく調査をもとにして「政治的権利」と「市民の自由」の2要因を測定している。「政治的権利」については公正で自由な選挙によって選ばれた政府の長および公正な選挙法など8項目から評価し、「市民の自由度」については言論の自由、結社の自由、法の支配、人権、経済的権利など14項目から評価している(近藤 2003)。指標は1972年以来毎年公表されており、ガバナンスと経済発展の関係に関

する経済学分野の研究ではガバナンス指標と して一般的に用いられている。指標はガバナ ンスの中の民主化という側面を捉えたもの で、多年度にわたって測られている。

- ・トランスペアレンシー・インターナショナル: 汚職実感指標 (CPI: Corruption Perceptions Index) 178ヵ国 (Transparency International 2010)を対象としており、民間の企業家、研究者、内部・外部者等による主観的投票による各種の調査を統合した指標で、政治家と行政の汚職の度合いに焦点を当てている。指標は1995年以来毎年公表されており、汚職関連の指標としては最も一般的で、投資家や政策決定者に利用されるだけでなく、汚職問題を公にして、この問題が多い国への外圧となる効果もある(近藤 2003)。指標はガバナンスの中の汚職という側面を捉えたもので、多年度にわたって測られている。
- ・PRSグループ:国際カントリー・リスク・ガイド (ICRG International Country Risk Guide)

140ヵ国を対象としており、評価は外部者の主観的な格付けになっている。指標は政治リスク、財政リスク、経済リスクの3分野に分かれており、政治リスクは「政府の安定性」「社会経済条件」「投資状況」「内部抗争」など12の要因、財政リスクは「対外債対国内総生産比率」「経常収支対輸出比率」など5要因、経済リスクは「GDP/cap」「年間インフレ率」など5要因が測定されている。指標は1980年以来毎月公表されており、企業機関に重視されている(近藤 2003)。指標はガバナンスの様々な側面を捉えており、多年度にわたって測られているもので、基本的に世界各国に投資している企業の需要に合うように作成されているため、内容が経済に特化している。

・世界銀行:国別政策・制度評価(CPIA: Country Policy and Institutional Assessment)

国の政策・制度の枠組みが、持続可能な成長 と貧困政策、開発援助の効果的な活用をどの 程度促進しているかを測っており、指標は経 済運営、構造政策、社会的一体性・公平のた めの政策、公共セクターの管理・制度の4分 野に分かれている。経済運営は「マクロ経済 運営」「財政政策」などの3要因、構造政策は「貿易」「金融セクター」など3要因、社会的一体性・公平のための政策は「ジェンダー平等」「公的資金利用の公平性」など5要因、公共セクターの管理・制度は「所有権とルールに基づくガバナンス」「予算・財務管理の質」などの5要因から構成されている(国際協力機構 2008)。指標はガバナンスの様々な側面を捉えており、情報は2000年より部分的に公開され、2006年からは全面的に毎年公開されている。CPIAの結果は世界銀行の国別資金配分を決定するために用いられている。

・世界銀行研究所:世界ガバナンス指標(WGI: Worldwide Governance Indicators)

212ヵ国・地域(Kaufmann, D. et al. 2009)を対象としており、政府等の公的機関、民間企業、NGOの専門家、一般市民や企業への調査などのデータを基にした440余の質問項目を統計処理し、指標を作成している。指標は「国民の声(発言力)と説明責任」「政治的安定と暴力の不在」「政府の有効性」「規制の質」「法の支配」「汚職の抑制」の6要因から構成されている(国際協力機構 2008)。指標は1996年以来隔年に公表され、2002年からは毎年の公表になっている。指標はガバナンスの様々な側面を捉えており、多年度にわたって測られている。

このようにガバナンス指標にはそれぞれの特徴があるので、援助効果との関連でガバナンスの何をみるかによって、目的にあったガバナンス指標を用いることが必要である。

## 2. 研究目的

国際協力機構は2003年から2008年にわたって技術協力プロジェクト終了時評価の2次評価を行っている。2次評価とは国際協力機構内部者が行ったプロジェクトの終了時評価について、国際協力機構外部の者が2次的にそのプロジェクトを評価することである。評価とは事実関係を確認して、その是非について判断をくだすことである。終了

時評価報告書には事実関係と1次評価者の判断が 記載されているが、2次評価者は報告書に記述さ れた1次評価者の判断部分ではなく事実の部分を 基にプロジェクトを判断する。報告書に事実関係 についての確認が十分にされずに判断が記述され ている場合、読み手としてはその内容を信じるか どうかということになるが、2次評価者は事実を 的確に確認することによって、その事柄に対して 2次評価者が報告書に記述されている判断と異な る判断をすることができるのではないかと考えら れる。評価者が一人であると、くだされた判断に その評価者の傾向が現れることは否めない。そこ で2次評価では評価をする際の判断基準などにつ いて評価者が統一の見解を共有するよう、前もっ ていくつかの評価例を用いてワークショップを行 った後、数人の評価者が1つの報告書を読み、評 価を行う。多くの人が評価に加わることによって、 偏りの少ない、信頼のおける結論が導き出される ことが期待できる。さらに、評価者達の評価結果 の平均をとることによって評価者の評価傾向は薄 まることになる。また、評価された値は統計分析 の仕方を工夫することによって評価者の評価傾向 と真の評価値とに分けられ、推定された真の値の みを評価結果として用いている。

このような評価手段・方法を用いた2次評価は、1次評価の結果をもとに評価を行うという制約はあるが、終了時評価を行った1次評価者とは異なる10人の2次評価者が同じプロジェクトを評価することによって、終了時評価の客観性を担保している。2次評価によるプロジェクト評価はDAC(開発援助委員会)の評価5項目であるプロジェクトの妥当性、効率性、有効性、インパクト、自立発展性に関して数量的評価を行っている。

2006年度の2次評価ではプロジェクト効果の発現に影響を与える要因を明らかにすべく、「プロジェクト規模」「プロジェクトの協力期間」、被援助国の「GDP/cap」「教育指数」「人間開発指数」の5要因と評価5項目に関する視点および評点項目との相関分析を行っている(国際協力機構 2006)。その結果、これら5つの影響要因とプロジェクトの「妥当性」「効率性」「有効性」「インパクト」に関する評価項目とには統計的に有意な相関はほとんどみられなかった。一方、「自立発展性」に

関する視点・評点項目と「GDP/cap」「教育指数」 「人間開発指数」との間に有意な相関がみられた。

これらの結果は発現したプロジェクト効果を活用し、持続させるためには被援助国の努力や仕組みといったガバナンス、経済力、教育という外部条件の影響を受けることを示唆している。しかし、この分析は相関分析であることから、被援助国側の要因が自立発展性に影響を及ぼすという因果関係を示しているのではない。

技術移転を主とした援助プロジェクトは、プロジェクト終了後に被援助国自身が移転された技術を活用し、独自にプロジェクトを実施していくことができるようにすることを目的としており、プロジェクト効果の発現や持続発展に対してその国のガバナンス状況がどのような影響を及ぼすかを明らかにすることは重要である。しかし、日本ではODAプロジェクトの効果に対する被援助国のガバナンスの影響に関する定量的分析はこれまでされていない。

プロジェクトの効果はプロジェクト開始時のインプットからプロセスを経て発現することから、プロジェクトの実施計画、実施過程が効果の発現に影響を与える。それゆえ、単にガバナンスとプロジェクト効果の発現あるいは効果の持続性との関係をみるのではなく、プロジェクトのインプットからアウトプット、アウトカムに至る一連の流れに位置するプロジェクト効果とガバナンスとの因果関係、また、効果の持続性とガバナンスとの因果関係を構造的に捉えることが重要である。

本研究では、国際協力機構が実施している技術協力プロジェクトの効果の発現(有効性)、および効果の活用と持続性(自立発展性)に対する被援助国のガバナンス、経済力、教育の影響について明らかにする。まず、ガバナンスと有効性および自立発展性との関連性を明らかにし、次に、プロジェクトのインプットから有効性、自立発展性に至る因果関係、被援助国のガバナンスから有効性、自立発展性に至る因果関係を想定して、プロジェクト効果の発現と効果の活用・持続性に及ぼすガバナンスの影響を構造的に明らかにする。さらに、良いガバナンスは経済発展や教育向上へ影響を与える(World Bank 1998)ことから、被援助国のガバナンスが経済や教育に影響を及ぼし、

その経済や教育が有効性および自立発展性に影響を及ぼすという因果関係を想定し、ガバナンス、経済力、教育程度がプロジェクトの有効性および自立発展性に及ぼす影響を明らかにする。

## 3. 研究の方法

## (1) 分析対象

分析には国際協力機構が2005年度から2008年度 に行った技術協力プロジェクト終了時評価の2次 評価によるデータを用いた(国際協力機構 2005、 国際協力機構 2006、国際協力機構 2007、国際協 力機構/日本評価学会 2009)。

国際協力機構では技術協力プロジェクト(プロジェクト)を多地域にわたり多くの国で実施しており、実施した案件全てについて終了時評価を行っている(国際協力機構2011a)。2003年度から2007年度までに行われた終了時評価実施案件数(フェーズ2、フェーズ3を除く)は、事業評価年次報告書(国際協力機構2005~2007)および国際協力機構(2011b、2011c)によると321案件である。本研究で分析対象としたプロジェクトは、2003年度から2007年度に終了した案件のうち2次評価を行ったもので、フェーズ2およびフェーズ3

表1 終了時評価実施案件数と分析対象数

	終了時評価	写施案件	分析対	象案件
	案件数	%	案件数	%
【地域】				
アジア大洋州	168	52.3	95	52.5
中南米	68	21.2	40	22.1
アフリカ	52	16.2	30	16.6
中近東	21	6.5	9	5.0
東欧	12	3.7	7	3.9
計	321	100.0	181	100.0
【分野】				
社会開発	137	42.7	77	42.5
農業開発	64	19.9	35	19.3
保健医療	54	16.8	33	18.2
森林自然	44	13.7	23	12.7
鉱工業開発	22	6.9	13	7.2
計	321	100.0	181	100.0

(出所) 筆者作成

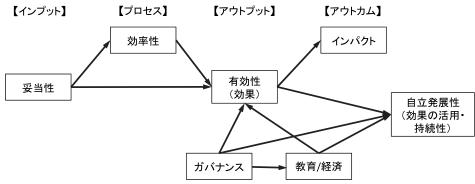


図1 想定図:プロジェクトの有効性、自立発展性へ影響を与える要因の因果関係

(出所) 筆者作成

の案件を除外した181案件である。表1は終了時評 価実施案件数と分析対象案件数を地域別・分野別 に示したものである。終了時評価実施案件と分析 対象案件の分布のバラツキを比較すると、両者間 の各地域・分野に占める割合にはいずれも統計的 に有意な差はみられず、分析対象案件は終了時評 価が実施された案件を代表しているといえる。

なお、2次評価はDACの評価5項目であるプロジェクトの妥当性、効率性、有効性、インパクト、自立発展性の各評価項目について5段階のリッカート・スケールで評価している。また、国際協力機構では終了時評価をプロジェクト終了の数ヶ月前に実施している。そのため、有効性、インパクト、自立発展性については終了時の状況の可能性を含めた評価である。

### (2) 分析方法

## ①プロジェクトのインプットからアウトカムに至る因果関係の想定

プロジェクトのインプット(妥当性)からプロセス(効率性)、アウトプット(有効性)、アウトカム(インパクト)、自立発展性に至る因果関係、および被援助国のガバナンス、教育/経済から有効性、自立発展性に至る因果関係を想定する。図1は想定した因果関係を示したものである。

## ②プロジェクトの妥当性、効率性、有効性、自立 発展性に関する観測変数

「妥当性」の概念変数に関して2評価項目、「効

率性」の概念変数に関して2評価項目、「有効性」 の概念変数に関して2評価項目、「インパクト」の 概念変数に関して2評価項目、「自立発展性」の概 念変数に関して4評価項目を観測変数として用い た (表2)。

### ③ガバナンス指標

分析対象のプロジェクトは実施開始時期が1998 年度~2002年度、終了時期が2003年度~2007年度 にわたっているため、分析に用いるガバナンス指 標は1998年から継続的に測定された指標であるこ とが必要である。また、当該国のガバナンスの一 側面からの影響ではなく、ガバナンス全体の状況 からの影響を明らかにすることを目的としている ことから、分析には共通の基準で多年度にわたっ た推定値が示されており、ガバナンス状況を様々 な側面から捉えて、大枠で把握できる世界銀行研 究所の世界ガバナンス指標(Kaufmann, et al 2009) を用いた (表3)。2009年度版では1996年、1998年、 2000年、2002年~2008年までの6つの各要因の推 定値が同じ基準のもとに算出されている。推定値 は-2.5から2.5の値をとり、-2.5は状況が最も悪く、 2.5は状況が最も良いことを示している。ガバナ ンスの影響を総体的に捉えるために、これら6要 因の推定値から主成分得点を算出して「ガバナン ス | の概念として用いた。

## 4分析手順

分析1では単回帰分析を用いて「ガバナンス」

## 表2 概念を構成する評価項目

### 【妥当性】

- ・必要性:プロジェクト実施の必要性はあるか
- ・手段としての適切性:プロジェクトの設計は正しかったか

### 【効率性】

- ・費用対効果:より安価な費用でより多くの効果を達成するよう工夫がなされていたか
- ・実施フプロセスの適切性:実施プロセスは適切だったか

### 【有効性】

- ・プロジェクト目標の達成度:プロジェクトは達成されているか
- ・プロジェクト目標達成に係わる因果関係の論理性:プロジェクト目標達成に至るまでの 因果関係は強いか

## 【インパクト】

- ・予期されたインパクトの達成度:プロジェクト目標達成により計画した効果は達成されたか
- ・インパクトに係わる因果関係の論理性:プロジェクト目標達成により予想される効果に 至るまでの因果関係は強いか

### 【自立発展性】

- ・持続性の程度:プロジェクトが目指した効果は援助終了後も持続しているか
- ・持続性を保証する環境の整備度(組織):持続性が保たれる組織力があるか
- ・持続性を保証する環境の整備度(技術):持続性が保たれる技術力があるか
- ・持続性を保証する環境の整備度(財政):持続性が保たれる財政力があるか

(出所) 筆者作成

### 表3 世界銀行研究所による世界ガバナンス指標

6つの要因	要因の意味
国民の声(発言力)と説明責任 (Voice and Accountability)	国民の政治参加(自由かつ公正な選挙など)、結社の自由、報道の自由があるかどうか。
政治的安定と暴力の不在 (Political Stability and Absence of Violence)	国内で発生する暴動(民族間の対立を含む)やテロリズムなど、制度化されていない、あるいは暴力的な手段により政府の安定が揺るがされたり、転覆される可能性がどれだけあるか。
政府の有効性 (Government Effectiveness)	行政サービスの質、政治的圧力からの自立度合い、政府による政策策定・実施への信頼度、政府による(改革への)コミットメント。
規制の質 (Regulatory Quality)	その国の政府が民間セクター開発を促進するような政策や規制を策定し、それ を実施する能力があるかどうか。
法の支配 (Rule of Law)	公共政策に携わる者が社会の法にどれだけ信頼をおいて順守しているか。特に契約の履行、警察・裁判所の質や犯罪の可能性など。
汚職の抑制 (Control of Corruption)	その国の権威・権力が一部の個人的な利益のために行使される度合い。汚職 の形は大小を問わず、また一握9のエリートや個人の利害関係による支配も含 む。

(出所) 国際協力機構 (2008) より作成

とプロジェクトの「有効性」および「自立発展性」 との関連性を明らかにする。

分析2では共分散構造分析法を用いて因果モデルを構築し、プロジェクトのインプットからアウトカムまでの流れの中から「ガバナンス」がプロジェクトの「有効性」および「自立発展性」に及

ぼす影響を構造的に捉え、因果関係を明らかにする。

分析3では「ガバナンス」と「教育」、「経済」 がプロジェクトの「有効性」および「自立発展性」 に及ぼす影響を、共分散構造分析を用いて明らか にする。なお、分析にはSPSS社のAmos™16.0を 用いた。

## 4. 結果

## (1) 評価項目について記述統計結果

表4はプロジェクトに対する評価点の平均を示したものである。「妥当性」については5つの評価項目のなかで相対的に高い評価を得ており、プロジェクト実施の必要性は高く、用いる手段は適切であると評価されている。しかし、手段の適切性に関する評価はバラツキが相対的に大きい。「効率性」については、得点が3点に達しておらず普通以下と評価されており、その中でも費用対効果に関する評価は低い。「有効性」については相対的に高く評価されており、プロジェクトの目標は比較的達成されており、目標達成に至る因果関係がみられる。「インパクト」に対する評価は相対的に低く、想定されたインパクトの達成度は普通以下で、評価のバラツキが相対的に大きい。「自

立発展性」についての評価は相対的に高く、プロジェクト終了後に効果を持続させるための技術や組織力をある程度持ち、持続可能であると評価されている。しかし、財政力については普通の水準以下と評価されている。組織力や財政力の評価はバラツキが大きく、プロジェクトによる差が大きい。

## (2) 分析1: ガバナンスとプロジェクトの有効性 および自立発展性との関連性

プロジェクトが開始されてから終了するまでの間、被援助国のガバナンス状況は変化している。国の政策などが実施され、効力がでてくるまでには時間がかかることから、「ガバナンス」の状況がプロジェクトの「有効性」や「自立発展性」に対して同時進行で反映されるのではなく、時間をある程度経てから反映されるというタイムラグが生じると考えられる。そこで、各プロジェクトの終了時期の1年前から5年前までの各年のガバナンスに関する6要因の推定値から各年の「ガバナンス」を算出した。なお、1999年度については1998

平均値 最大値 最小値 標準偏差 【妥当性】 3.74 4.45 2.61 0.363 必要性 手段としての適切性 3.21 4.27 1.73 0.462 【効率性】 2.77 3.92 2.00 0.374 費用対効果 実施プロセスの適切性 2.91 4.02 1.82 0.401 【有効性】 4.12 2.00 0.438 プロジェクト目標の達成度 3.16 3.37 3.98 2.42 0.302 因果関係の論理性 【インパクト】 インパクトの達成度 2.66 3.81 1.13 0.491 因果関係の論理性 2.87 3.82 1.63 0.412 【自立発展性】 4.18 1.94 0.466 持続性の程度 3.15 3.09 4.44 1.91 0.526 持続性を保証する組織力 4.07 2.01 0.421 持続性を保証する技術 3.14 持続性を保証する財政 2.68 4.12 1.52 0.514

表4 プロジェクトに対する評価得点

評価基準 1.不十分/低い ~ 5.十分/高い

(出所) 筆者作成

	自立発展性					
ガバナンス	偏回帰係数	標準誤差	t値	$R^2$	調整済R <sup>2</sup>	F値
終了時1年前	0.195	0.073	2.665	0.038	0.033	7.104
終了時2年前	0.191	0.073	2.608	0.037	0.031	6.802
終了時3年前	0.169	0.074	2.296	0.029	0.023	5.273
終了時4年前	0.197	0.073	2.687	0.039	0.033	7.223
終了時5年前	0.187	0.073	2.545	0.035	0.030	6.480

表5 ガバナンスと自立発展性の関連性

(出所) 筆者作成

年度、2001年度については2000年度の推定値を用いた。

「ガバナンス」と「有効性」との関係、および「ガバナンス」と「自立発展性」との関連をみるために、各年の「ガバナンス」を説明変数、「有効性」および「自立発展性」をそれぞれ被説明変数として単回帰分析を行った。「有効性」、「自立発展性」は各々の評価項目を用いた主成分得点による合成変数である。

「ガバナンス」と「有効性」との各時期における回帰モデルはいずれも有意ではなく、「ガバナンス」と「有効性」との間に関連性はみられなかった。

一方、「自立発展性」に関しては、「ガバナンス」はいずれの時期においも「自立発展性」と1%または5%水準で有意な関連性がみられる(表5)。これは、「ガバナンス」状況が良くなるほど、被援助国においてプロジェクトの効果が活用され、持続していくことを意味している。例えば、終了時1年前の「ガバナンス」状況とプロジェクトの「自立発展性」の関連をみると、プロジェクト終了時1年前の「ガバナンス」の状況が1ポイント良くなれば、その状況が1年後に反映されてプロジェクト終了後の効果の活用や持続性の程度が0.195高くなる、ということである。

これらの結果から「ガバナンス」はプロジェクト効果の発現「有効性」との関連性はみられないが、いずれの時期においても「ガバナンス」と「自立発展性」とには関連性があることが明らかになった。

しかし、効果の発現にはプロジェクト実施における手段の適切性や実施プロセスにおける効率性

などが影響を及ぼすと考えられる。さらに、「自立発展性」は被援助国側が発現した効果を活用し、持続させることができるようになることを意味しているので、それにはまず効果が発現することが大事である。それゆえ、「有効性」から「自立発展性」へと高い関連性があることが想定される。

そこで、分析2では、プロジェクトの「妥当性」から「効率性」、「有効性」、「自立発展性」に至る関係、および「有効性」、「自立発展性」に対する「ガバナンス」の関係を構造的に捉えて、「ガバナンス」の影響を明らかにする。

## (3) 分析2: ガバナンスがプロジェクトの有効性 および自立発展性に及ぼす影響

### (1) 「ガバナンス |

各時期の「ガバナンス」を用いて因果モデルを構築し、どの時期の「ガバナンス」がモデルとして適切であるかについて、モデルの適合度を比較した。複数のモデルを比較する場合に用いられるAIC(赤池の情報量基準)指標は値が小さいほどそのモデルが優れていると判断される指標であるが、AICの値が最も小さいのはプログラム終了時5年前の「ガバナンス」であった(表6)。

この結果から、分析には「ガバナンス」の状況を示す変数として、各プログラム終了時5年前の推定値による「ガバナンス」を用いた。なお、「ガバナンス」は主成分得点による合成変数であるが、この変数におけるガバナンス指標に関する6要因の成分は表7に示すとおりで、「ガバナンス」全体の73.5%を説明している。「ガバナンス」に対して「法の支配」が最も寄与しており、次いで「政府の有効性」、「汚職の抑制」、「規制の質」と

表6 年度による因果モデルの比較

因果モデル	GFI	AGFI	RM SEA	AIC
終了時1年前	0.929	0.891	0.055	155.01
終了時2年前	0.930	0.892	0.054	154.01
終了時3年前	0.931	0.893	0.053	152.93
終了時4年前	0.931	0.894	0.053	152.60
終了時5年前	0.931	0.894	0.052	151.72

(出所) 筆者作成

表7 ガバナンスの成分

6つの要因	第1主成分(ガバナンス)
法の支配	0.947
政府の有効性	0.934
汚職の抑制	0.905
規制の質	0.879
政治的安定と暴力の不在	0.755
国民の声(発言力)と説明責任	0.692

(出所) 筆者作成

なる。「国民の声(発言力)と説明責任」と「政 治的安定と暴力の不在」は「ガバナンス」に対す る寄与が相対的に低く、その中でも「国民の声 (発言力)と説明責任」は「ガバナンス」に対す る寄与が最も低い。この「ガバナンス」変数は社 会の法への信頼や政府の政策実行との関連が強 く、民主化の側面をもつ国民の自由や政府の安定 との関連は相対的に低いと捉えることができる。

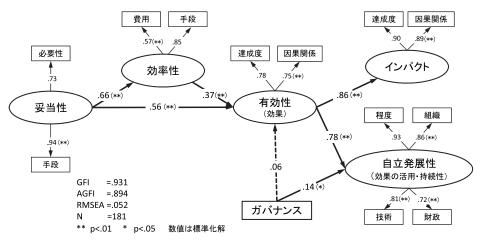
## ②ガバナンスの影響

図2は共分散構造分析によって推定されたプロジェクトの「妥当性」から「自立発展性」に至る因果関係、および「ガバナンス」から「有効性」、「自立発展性」に至る因果関係を示してある。

図中、実線は有意な関係、点線は有意ではない関係を示し、観測変数や潜在変数に関連した誤差変数の表示は省略してある(図3、図4も同様)。モデルの全体的評価として、GFI(適合度指標)、AGFI(修正済適合度指標)、RMSEA(平均二乗誤差平方根)の指標を用いた。このモデルの適合度はGFI=0.931、AGFI=0.894、RMSEA=0.052であった。GFIが0.9以上、RMSEAが0.05以下であるとモデルの適合度が高いと判断できる(豊田2007)。このモデルのGFIは0.9以上あり、RMSEAは0.052であるので適合度は高く、データを良く説明していると判断できる。また、構成概念から観測変数への影響指標の値はいずれもワルド検定の結果1%水準で有意となっており、構成概念と観測変数の関係は適切に対応しているといえる。

まず、構成概念間の因果係数によって想定した 因果関係について述べる。「妥当性」から「有効 性」への係数値は0.564で、1%水準で有意な因果 関係がみられる。プロジェクト実施に際し、適切

図2 ガバナンスがプロジェクトの有効性、自立発展性に及ぼす影響



(出所) 筆者作成

な手段をとるなどプロジェクトの妥当性が高いとプロジェクトの効果が発現しやすいことを示している。また、「妥当性」から「効率性」の係数値は0.662で、1%水準で有意な因果関係がみられ、プロジェクトの妥当性が高いと実施プロセスにおける効率性が高まることを示している。さらに、「効率性」から「有効性」への係数値は0.365で、1%水準で有意な因果関係があり、実施プロセスにおける効率性が高いとプロジェクトの効果が発現しやすいといえる。

「有効性」から「インパクト」への係数値は 0.862で、1%水準で有意な因果関係がみられ、プロジェクトの効果が多く発現するとインパクトが 生じやすいことを示している。また、「有効性」から「自立発展性」への係数値は0.780で、1%水準で有意な因果関係がみられ、プロジェクト効果が多く発現すると、その成果を活用し、持続させようとする「自立発展性」が高くなることを示している。

この結果から、「妥当性」は「効率性」と「有効性」に対して影響を及ぼし、「効率性」は「妥当性」からの影響をうけると同時に「有効性」へ影響を及ぼし、さらに、「有効性」は「妥当性」と「効率性」からの影響を受け、「インパクト」と「自立発展性」に対して影響を及ぼすという因果関係があるといえる。

次に、「ガバナンス」が「有効性」および「自立発展性」に及ぼす影響をみていく。「ガバナンス」から「有効性」への係数値は0.057で有意な因果関係はみられず、「ガバナンス」は「有効性」に対してあまり影響を及ぼしていない。一方、「ガバナンス」から「自立発展性」への係数値は0.136で、5%水準で有意な因果関係がみられ、ガバナンス状況が良いと「自立発展性」が促進されることを示している。「ガバナンス」の「自立発展性」に対する影響をみると、「自立発展性」への直接的影響と「有効性」を経た間接的影響があり、合計すると「ガバナンス」から「自立発展性」に対する影響は0.180となる。

この結果から「ガバナンス」はプロジェクト効果の発現に対してはあまり影響を及ぼさないが、発現した効果を活用し、持続させるという「自立発展性」に対しては影響を及ぼすと考えられる。

「自立発展性」は被援助国の「ガバナンス」からの影響をうけるとともに、「有効性」からも影響をうけている。「ガバナンス」からの影響力と「有効性」からの影響力を比べると、「有効性」からうける影響力の方が大きい。

## (4) 分析3: ガバナンス、教育、経済がプロジェ クトの有効性および自立発展性に及ぼす影響

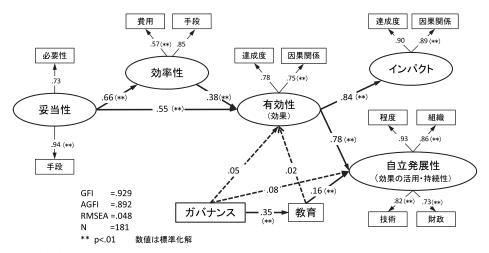
被援助国の教育の程度、経済力を表す指標としては国連開発計画による「人間開発報告書」(国連開発計画 2000年~2008年)に記載されている教育指数、GDP/capを用い、「教育」と「経済」の変数とした。教育指数は成人識字率、初等・中等・高等教育への就学率を総合した指数である。GDP/capは自然対数値に変換して用いた。ガバナンス状況を表す変数として用いたのは分析2で用いた「ガバナンス」である。また、「教育」および「経済」変数は、当該国の教育程度、経済力に「ガバナンス」が反映されるには時間がかかると想定し、因果モデルを構築して比較した結果、終了時1年前の推定値を用いたモデルのAIC値が最も小さいため、各プロジェクト終了時1年前の教育指数、GDP/capの値を用いた。

## ①ガバナンスと教育がプロジェクトの有効性および自立発展性に及ぼす影響

図3はプロジェクトの「妥当性」から「自立発展性」に至る因果関係、「ガバナンス」から「有効性」、「自立発展性」に至る因果関係、および「ガバナンス」が「教育」を経て「有効性」、「自立発展性」に至る因果関係を示したものである。このモデルの適合度はGFI=0.929、AGFI=0.892、RMSEA=0.048であるので、適合度は高く、データを良く説明している。

「妥当性」から「自立発展性」に至る因果関係をみると、「妥当性」から「有効性」への係数値、「妥当性」から「効率性」への係数値、「効率性」から「有効性」への係数値、「有効性」から「自立発展性」への係数値、さらに「有効性」から「インパクト」への係数値はいずれも1%水準で有意であることから「妥当性」から「自立発展性」、「インパクト」への因果関係は成立している。

「ガバナンス」から「有効性」、「自立発展性」



## 図3 ガバナンスと教育がプロジェクトの有効性および自立発展性に及ぼす影響

(出所) 筆者作成

への直接的な関係を示す係数値は0.051と0.078で有意な因果関係はみられない。「教育」ついては、「ガバナンス」から「教育」への係数値が0.354で1%水準で有意な因果関係がみられる。さらに、「教育」から「自立発展性」への関係については、係数値が0.162で1%水準で有意な因果関係がみられる。一方、「教育」から「有効性」への係数値は0.016で、有意な因果関係はみられない。「ガバナンス」の「自立発展性」に対する影響をみると、「自立発展性」への直接的な影響と「有効性」や「教育」を経た間接的な影響があり、合計すると「ガバナンス」から「自立発展性」に対する影響は0.180となる。

これらの結果から、「ガバナンス」は「有効性」、「自立発展性」に対して直接的にはあまり影響を及ぼさないが、「自立発展性」に対しては「教育」を介して影響を及ぼしていると考えられ、「ガバナンス」状況が良いとその国の教育程度が向上し、プロジェクトの効果の活用・持続性が促進される、と考えられる。「自立発展性」は「教育」と「有効性」からの影響をうけており、「教育」よりも「有効性」からうける影響力の方が大きい。

## ②ガバナンスと経済がプロジェクトの有効性および自立発展性に及ぼす影響

図4はプロジェクトの「妥当性」から「自立発

展性」に至る因果関係、「ガバナンス」から「有効性」、「自立発展性」に至る因果関係、および「ガバナンス」が「経済」を経て「有効性」、「自立発展性」に至る因果関係を示したものである。このモデルの適合度はGFI=0.930、AGFI=0.894、RMSEA=0.046であるので、適合度は高く、データを良く説明している。

「妥当性」から「自立発展性」に至る因果関係をみると、「妥当性」から「有効性」への係数値、「妥当性」から「効率性」への係数値、「効率性」から「有効性」への係数値、「有効性」から「自立発展性」への係数値、さらに「有効性」から「インパクト」への係数値はいずれも1%水準で有意であることから「妥当性」から「自立発展性」、「インパクト」に至る因果関係は成立している。

「ガバナンス」から「有効性」、「自立発展性」への係数値は0.094と0.029で、有意な因果関係はみられない。「経済」については、「ガバナンス」から「経済」への係数値は0.638で、1%水準で有意な因果関係がみられる。さらに、「経済」から「自立発展性」への関係をみると、係数値は0.171で1%水準で有意な因果関係がみられる。一方、「経済」から「有効性」への係数値は-0.060で有意な因果関係はみられない。「ガバナンス」が「自立発展性」に対して及ぼす影響には直接的な影響と「有効性」や「経済」を経た間接的な影響

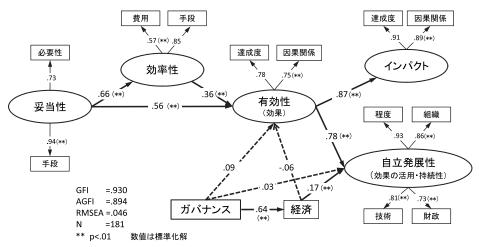


図4 ガバナンスと経済がプロジェクトの有効性および自立発展性に及ぼす影響

(出所) 筆者作成

があり、これらを合計すると「ガバナンス」から 「自立発展性」に対する影響は0.181となる。

これらの結果から、「ガバナンス」は「有効性」、「自立発展性」に対して直接的にはあまり影響を及ぼさないが、「自立発展性」に対しては「経済」を介して影響を及ぼしていると考えられ、「ガバナンス」状況が良いとその国の経済が発展し、プロジェクト終了後におけるプロジェクト効果の活用・持続性が促進される、と考えられる。「自立発展性」は「経済」と「有効性」から影響をうけており、「経済」よりも「有効性」からうける影響力の方が大きい。

## 5. 結論と考察

本研究は、被援助国側の「ガバナンス」が日本の技術協力プロジェクトの効果の発現、および発現した効果の活用・持続性に与える影響を明らかにすることを目的としている。

回帰分析法によって、「ガバナンス」はプロジェクト効果の発現「有効性」に対してはあまり関連性はみられないが、発現した効果の活用・持続性に関する「自立発展性」と「ガバナンス」には高い関連性があることを明らかにした。さらに、プロジェクトを総体的に捉えた場合の「有効性」

及び「自立発展性」に対する「ガバナンス」の影 響を明らかにするために因果モデルを構築し、検 証した。分析の結果、想定したモデルとデータの 適合度は高く、構成概念間の因果関係を検討する ことによって「ガバナンス|が「有効性|にはあ まり影響を及ぼさないが、「自立発展性」に対し ては影響を及ぼすことを明らかにした。また、被 援助国のガバナンス、教育程度および経済力の関 連から「有効性」、「自立発展性」に対する影響を みると、「ガバナンス」は「自立発展性」に対し て影響を直接に及ぼすのではなく、「教育 | や 「経済」をとおして間接的に影響を及ぼすことが 示された。「自立発展性」は「有効性」と「ガバ ナンス」、あるいは「有効性」と「教育」、「有効 性 | と「経済 | から影響をうけているが、「自立 発展性」がうける影響の強さを比べると、いずれ の場合も「有効性」からの影響力の方が強い。

これらの結果から、プロジェクト効果の発現という「有効性」が「自立発展性」に対して影響を与えるばかりでなく、被援助国のガバナンス状況、さらにガバナンスの影響をうけた経済力、教育程度が「自立発展性」に対して影響を与えていることが明らかになった。

援助のアウトカムとして識字率の高まり、乳児 死亡率の低下を示す研究がある(World Bank 1998)。このようなアウトカムは援助終了後すぐ に発現するものではなく、アウトプットがある期間継続してから発現するのであるから被援助国がそのアウトプットを持続させることが必要である。ODA政策の基本方針の一つとして、途上国の自助努力を支援することをあげている。援助国側が援助効果の発現を図ることは当然のことであり、効果の発現は妥当性や効率性との因果関係が強いことに鑑み、被援助国の自立発展の支援を図るために、プロジェクトの実施計画の作成段階で援助国側が被援助国のガバナンス状況やキャパシティーの程度に留意した上で、手段などを工夫し、実施することが重要であると考えられる。

## 6. 今後の課題

本研究が分析の対象とした技術協力プロジェク トは技術・知識の移転を行いながら、一定期間事 業運営に関する協力を行い、協力終了後は被援助 国の運営に引き継がれていくものである。分析に 用いたデータは終了時評価報告書に基づいたデー タであることから、既述したように終了時評価実 施時点の有効性、インパクト、自立発展性につい ての評価はプロジェクト終了時における可能性を 含めたものである。今後、事後評価報告書のデー タなどを用いて、予想どおりにプロジェクトの効 果・インパクトが発現しているか、また、被援助 国が発現した効果を活用し、自立発展できている かを確認できた時点でのデータ分析を行うことに よって、プロジェクトの有効性・インパクト・自 立発展性に対するガバナンスの影響を明確にする ことが必要であろう。さらに、本研究ではデータ の制約上プロジェクトの分野別・地域別あるいは 内容別の分析はできなかった。しかし、ガバナン スと自立発展性との相関を分野別でみるといずれ も相関はみられなかったが、地域別にみると相関 がみられる地域と相関がみられない地域がある。 今後データを増やし、分野別・地域別・内容別な どによるガバナンスの影響を明らかすることが必 要であると考える。

### 謝辞

本論は国際協力機構が行った2006年度2次評価を基に2005年から2008年度のデータを用いて再分析したものである。関係各位に感謝の意を表する。なお、ここで述べた見解は筆者個人のものであり、国際協力機構の見解を示すものではない。

### 参考文献

稲田十一 (2006) 「第1章 ガバナンス論を巡る国際的 潮流」、下村恭民編著『アジアのガバナンス』、有斐 閣、7

外務省(2004)『ODA政府開発援助白書2003年版』、 179

国際協力機構(2005)『事業評価年次報告書2005』、 111-144

国際協力機構(2006)『事業評価年次報告書2006』、 101-136

国際協力機構(2007)『事業評価年次報告書2007』、79-114

国際協力機構(2008)『指標から国をみる-マクロ経済指標、貧困指標、ガバナンス指標の見方』、国際協力機構・国際協力総合研修所、117-152

国際協力機構/日本評価学会(2009)『外部有識者事業評価委員会による技術協力プロジェクト終了時評価の2次評価結果 総合報告書』、国際協力機構

国際協力機構(2011a)

http://www.jica.go.jp/activities/evaluation/middle-end.html(参照日2011/7/29)

国際協力機構(2011b)

http://www.jica.go.jp/activities/evaluation/middle-end.html(参照日2011/7/29)

国際協力機構(2011c)

http://www2.jica.go.jp/ja/evaluation/index.php?anken=&a rea1=&country1=&area2=&country2=&area3=&country 3=&field1=&field2=&field3=&tech\_ga%5B%5D=%E7%B5%82%E4%BA%86%E6%99%82%E8%A9%95%E4%BE%A1&start\_from=1997&start\_to=2007&search=%E6%A4%9C%E7%B4%A2 (参照日2011/7/30)

国連開発計画 (2001-2007/2008) 『人間開発報告書2001』 ~『人間開発報告書2007/2008』、国際協力出版会 近藤正規 (2003) 『ガバナンスと開発援助-主要ドナー

- の援助政策と指標構築の試み-』、国際協力事業団・ 国際協力総合研究所
- 下村恭民(1999)「ガバナンス、経済発展、援助-実 効ある政策論議への脱皮を求めて-」、『開発援助研 究』、5(4):59-73
- 豊田秀樹(2007)『共分散構造分析 [Amos編]』、東京 図書
- Burnside, C. and Dollar, D. (1998). Aid, the Incentive Regime, and Poverty Reduction, Policy Research Working Paper No.1937, Washington DC: World Bank.
- Burnside, C. and Dollar, D. (2000). Aid, Policies, and Growth. *The American Economic Review*, 90(4), 847-868.
- Burnside, C. and Dollar, D. (2004). Aid, Policies, and Growth: Reply. *The American Economic Review*, 94(3), 781-784.
- Dalgaard, C., Hansen, H. and Tarp, F. (2004). On the Empirics of Foreign Aid and Growth. *The Economic Journal*, 114(June), F191-F216.
- Freedom House (2009)
  - http://www.freedomhouse.org/uploads/fiw10/FIW\_2010\_

- Tables\_and\_Graphs.pdf (参照日2011/3/15)
- Kaufmann, D., Kraay, A. and Mastruzzi, M. (2009).
  Governance Matters VIII Aggregate and Individual Governance Indicators 1996-2008. Policy Research Working Paper No.4978, Washington DC: World Bank.
- Knack, D. and Rahman, A. (2007). Donor fragmentation and bureaucratic quality in aid recipients. *Journal of Development Economics*, 83 (2007), 176-197.
- Knack, S. (2001). Aid Dependence and the Quality of Governance: Cross-Country Empirical Tests. Southern Economic Journal, 68(2), 310-329.
- Rajan, R. and Subramanian, A. (2007). Does Aid Affect Governance? *The American Economic Review*, 97(2), 322-327.
- Transparency International (2010)
  http://www.transparency.org/policy\_research/surveys\_indices/cpi/2010/results(参照日2011/3/15)
- World Bank (1998). Assessing Aid: What Works, What Doesn't, and Why. Washington DC: Oxford University Press, 36-39.

(2011.8.17受理)

# The Influence of Governance on the Effectiveness of JICA Technical Cooperation Projects

Reiko Kikuta

Hiromitsu Muta

Tokyo Institute of Technology Graduate School of Decision Science and Technology kikuta.r.aa@m.titech.ac.jp

Tokyo Institute of Technology
Graduate School of Decision Science and Technology
muta@hum.titech.ac.jp

#### **Abstract**

Research on aid effectiveness has focused on the relationship between aid effectiveness and the governance of developing countries. However, no research shows the relationship between the effectiveness of Japanese ODA projects and governance measured by quantitative analysis. This study aims to clarify the influence of governance in developing countries on the effectiveness of JICA technical cooperation projects and on the sustainability of their effects. 181 projects intended to assist developing countries between 1998 and 2007 were examined. From regression analysis, it was found that the governance measured by Worldwide Governance Indicators did not have any significant explanatory power on project effectiveness, but was significantly related to the sustainability of project effects. The influence of the developing country's governance, economics and education on the effects of the project was also examined by covariance structure analysis. The study showed that governance, economics and education did not affect project effectiveness, but did promote the sustainability of project effects. It was also discosverd that governance indirectly influenced the sustainability of project effects through economics or education.

#### **Keywords**

governance indicators, technical cooperation project, aid effectiveness, covariance structure analysis

### 日本評価学会春季第8回全国大会 「国の予算と評価 | 開催の報告とお礼

2011年6月11日開催(於:政策研究大学院大学)の日本評価学会第8回全国大会には94名の方々にご出席いただき誠にありがとうございました。会員各位の日頃の研究や実践活動の報告を基に、評価研究者、実務者の間の経験、情報、知識を共有化する場として、皆様にとって有益な機会となりましたら幸甚に存じます。多くの方のご尽力により本大会を開催することができましたこと、心より感謝申し上げます。今後とも当学会の活動にご高配賜りますようどうぞ宜しくお願い申し上げます。

実行委員長 廣野良吉 (政策研究大学院大学/成蹊大学) プログラム委員長 牟田博光 (東京工業大学)

#### 2011年6月11日(土)

9:30-10:00	受 付				
午前の部 10:00-12:00	共通論題 1 多彩なインパクト評価手法 :最新の著作から学ぶ (佐々木亮) 1A	共通論題 2 国の予算と評価 (林 薫) 1B	<b>自由論題 1</b> 評価手法 (山谷清志) IC		
お昼休み 12:00-13:15		お昼休み (理事会12:05-13:00)			
午後の部 I 13:15-15:45	ラウンド・テーブル・ディスカッション 「国の予算と評価」 (廣野良吉) 想海樓ホール				
午後の部 II 16:00-18:00	共通論題 3 評価者倫理・スタンダード案に 関するワークショップ (三好皓一) 1A	自由論題 2 事例分析 (高千穂安長) 1B	自由論題 3 教育評価/学校評価 (橋本昭彦) IC		

# 日本評価学会春季第8回全国大会「国の予算と評価」 プログラム詳細

# 6月11日(土) 受付 9:30-10:00

	6月11日(土) 午前の部	ß 10:	00-1	2:0	0
共通論	題 1 多彩なインパクト評価手法:最新の著作から学ぶ	座長	佐々木	亮	(株)国際開発センター 1A
K1-1	『ランダム化比較試験 (RCT) の設計:ヒューマンサービス、社会科学領域における活用のために』概説		原田	隆之	目白大学
K1-2	『プログラムを成功に導くGTOの10ステップ:計画・実施・評価のための方法とツール』	0	池田	琴恵 満	お茶の水女子大学 国際基督教大学教育研究所
K1-3	『評価論理:評価学の基礎』に見るインパクト評価		佐々木	・売	(株)国際開発センター
K1-4	アウトカム志向の実務者向けプログラム評価		矢代	隆嗣	(株)アリエールマネジメント ソリューションズ
# \ヹ ≡෮					
共通論	題 2 国の予算と評価	座長	林	薫	文教大学 1B
	超 2 国の予算と評価  予算と評価の連動:日本と諸外国の現状と課題	座長	林田中	<b>薫</b> 秀明	文教大学 18 政策研究大学院大学
		座長			
K2-1	予算と評価の連動:日本と諸外国の現状と課題 自治体行政評価の予算への活用-三重県の事例を	座長	田中	秀明	政策研究大学院大学
K2-1 K2-2	予算と評価の連動:日本と諸外国の現状と課題 自治体行政評価の予算への活用 – 三重県の事例を 参考にして 公共財政管理と予算の評価	座長	梅田林	秀明	文教大字     政策研究大学院大学     (株)日本能率協会コンサルティング
K2-1 K2-2	予算と評価の連動:日本と諸外国の現状と課題 自治体行政評価の予算への活用 – 三重県の事例を 参考にして 公共財政管理と予算の評価		梅田林	秀明次郎	文教大字     政策研究大学院大学     (株) 日本能率協会コンサルティング     文教大学
K2-1 K2-2 K2-3 自由論	予算と評価の連動:日本と諸外国の現状と課題 自治体行政評価の予算への活用-三重県の事例を 参考にして 公共財政管理と予算の評価 題 1 評価手法		田中 梅田 林 山谷	秀明次郎薫	又教大字         政策研究大学院大学         (株)日本能率協会コンサルティング         文教大学         同志社大学    1C
K2-1 K2-2 K2-3 自由論 J1-1	予算と評価の連動:日本と諸外国の現状と課題 自治体行政評価の予算への活用 - 三重県の事例を 参考にして 公共財政管理と予算の評価 題 1 評価手法 地方議員の活動の評価方法開発のための基礎調査		田中 梅田 林 山谷	秀明 次郎 薫 <b>清志</b> 正美	又教大字         政策研究大学院大学         (株)日本能率協会コンサルティング         文教大学         同志社大学         東京大学大学院

# 6月11日(土) お昼休み 12:00-13:15 (理事会 12:05~13:00)

	6月11日(土) 午後の部Ⅰ 13:15-15:45						
ラウンド・テーブル・ディスカッション「国の予算と評価」 座長 廣野 良吉 成蹊大学 想海樓ホール							
RTD-1	国会予算分析局と政策評価学	上	野 宏	南山大学			
RTD-2	国の予算編成・執行過程の透明化・洗練化に向けて	高	崎 正有	三菱UFJリサーチ&コンサルテイング			
RTD-3	「政策予算分析局」と政策評価-政策評価なのか業績予算	山	谷 清志	同志社大学			
	なのか?						
RTD-4	ラウンドテーブル「予算と評価」途上国の公共財政管理	林	薫	文教大学			
	(Public Financial Management) から学ぶこと						

	6月11日(土) 午後の部	II 16	: 00-	-18:0	00	
共通論	題 3 評価者倫理・スタンダード案に関するワークショップ	座長	三好	皓一	立命館アジア太平洋大学	1A
		討論者			岩手県立大学	
		討論者			明治大学	
K3-1	「評価者倫理・スタンダード  行動指針 (案) 策定に向け			木英多島		
110 1	たいくつかの論点と考察-日本評価学会員を対象とした質		,	1.7072	111111111111111111111111111111111111111	
	問紙調査等から得られた知見より一					
K3-2	「評価者倫理・スタンダード  行動指針 (Codes) 案に関す		岩渕	祐二	政策プランナー	
	る基本的な考え方(前半)					
K3-3	「評価者倫理・スタンダード」行動指針(Codes)案に関す		服部	浩昌	青山学院大学大学院	
	る基本的な考え方(後半)と今後の検討課題					
自由論	題 2 事例分析	座長	高千種	恵 安長	玉川大学	1B
J2-1	技術協力プロジェクトの効果に及ぼすガバナンスの影響	0	菊田	怜子	東京工業大学大学院	
			牟田	博光	東京工業大学大学院	
J2-2	都道府県かつ産業毎の生産効率性と生産要素の余剰:日本 の製造業に関して		吉田	素教	大阪府立大学	
J2-3	主要ドナー比較調査-政策レベルのODA評価手法・体	0	佐々	木 亮	(株)国際開発センター	
	制一		菊池	美帆子	(株)国際開発センター	
			清水	正治	(株)国際開発センター	
			高杉	真奈	(株)国際開発センター	
				良吉	成蹊大学	
			-	眞文	国際基督教大学	
J2-4	欧州における社会投資収益率(SROI)の発展とその評価	0	伊藤	健	慶應義塾大学	
			玉村	雅敏	慶應義塾大学	
自由論	題 3 教育評価/学校評価	座長	橋本	昭彦	国立教育政策研究所	1C
J3-1	学校経営とその評価		石田	謙豪	元広島県尾道市立栗原北小学校	
J3-2	人材育成プログラムの評価研究 - 欧州委員会における		野呂	高樹	(財)未来工学研究所	
	Marie Curie ActionsのEx-post Impact Assessment studyを 事例として一					
J3-3	学校評価の活用方法-地域運営学校における学校運営		清水	洋一	東京都公立小学校	
J3-4	委員会の役割- (都合により取りやめ) 日本と非英語圏EUへの短期留学についての評価結果の		佐藤	由利子	東京工業大学	
	比較から					

# 共通論題セッション報告

#### 共通論題セッション1「多彩なインパクト評価手法:最新の著作から学ぶ|

座長 佐々木 亮 ((株)国際開発センター)

この1~2年の間にインパクト評価の理論や手法の解説を含む評価の専門書が日本で相次いで出版された。本セッションでは、それらの著書の著者・翻訳者自身に論じてもらい、フロア参加者との議論を通じて、日本にけるインパクト評価の普及と学会の役割に関して探った。

原田隆之会員(目白大学)から、『ランダム化比較試験(RCT)の設計:ヒューマンサービス、社会科学領域における活用のために』(トーガーソン&トーガーソン)について解説いただいた。同書は、アウトカム評価の「黄金律」であるランダム化比較試験(RCT)についての解説書である。医療・ヘルスケア領域のみならず、社会科学領域においてもRCTを積極的に活用していくことが、今後ますます重要になってくる。RCTは「何に効果があるのか」という問いに答え、適切な意思決定をするためのエビデンスを提供する最も頑健な方法であるばかりか、きわめて倫理的な方法でもある。ただし、社会科学領域においては、まだまだ障壁が多く残されているという指摘がなされた。

池田琴恵(お茶の水女子大学)・池田満(国際基督教大学教育研究所)の両氏からは、『プログラムを成功に導く GTO の 10 ステップ:計画・実施・評価のための方法とツール』(チンマン,イム,&ワンダーズマン)について解説いただいた。コミュニティ心理学と評価学における"評価"の発展における本書の位置づけを明らかにしたうえで、本書におけるインパクト評価の意味と役割を解説いただいた。さらに、本書の翻訳にあたって直面した、日本における今後の課題についても解説があった。

佐々木亮 ((株)国際開発センター) は、『評価論理:評価学の基礎』に見るインパクト評価について発表を行なった。因果関係の特定に関して定性・定量の全ての方法の背後に横たわる基本的論理である「一般的排除法」(スクリヴェン、2007) について解説した。また、インパクト分析(インパクト査定)とインパクト評価の違いに関する議論と、インパクト評価の課題を示す。最後に、評価が持つ本来の社会的機能や有効性を世の中の関係者に再認識してもらい、評価活動を通じた社会の改善に貢献することの重要性に関して指摘する。

矢代隆嗣氏 ((株)アリエールマネジメントソリューションズ)「アウトカム志向の実務者向けプログラム評価」では、『プログラム評価入門』(マイケルスミス)について解説いただいた。同書は、アウトカムの実現をめざす実務家が'評価'を業務に活用する目的、手順、手法を簡潔に説明している。ヒューマンサービス分野において、プログラム改善を目的とした評価に欠かせない質的調査の考え方と手法の説明が評価実践を動機付けるという解説があった。また、実務家が業務に評価を活かすには問題解決力が必要であるとの指摘も示唆に富む。

これらの報告に対して会場からは多数の質問・意見が出され、インパクト評価が徐々に普及してきたことが感じられた。インパクト評価の手法については、今後とも継続的に議論がなされる必要があろう。

#### 共通論題セッション2「国の予算と評価|

座長 林薫(文教大学)

本セッションは大会全体のテーマである「国の予算と評価」について、議論を深めるために企画された。 共通論題での議論が総論とすれば、各論という位置づけができるものである。

第一報告の田中会員による「予算と評価の連動:日本と諸外国の現状と課題」では業績予算を中心に予算と評価の連動に関する基本的事項を議論した。業績予算は、資源配分における意思決定の改善、プログラムや施策等のマネジメントの改善、議会や国民に対する政府の説明責任、政府活動の透明度の向上などの目的をもって、主要先進国で導入されつつある。主要国の改革の動向を概観しつつ、日本としては、評価制度にとどまらず、予算制度・予算編成プロセスの改革、ゲームのルールの見直し、意思決定のレベル(内閣・省庁・部局レベル)に応じた仕組み、財務・会計責任の明確化と議会の関与など、広範な課題に取り組む必要性があることが議論された。

第二報告の梅田会員による「自治体行政評価の予算への活用—三重県の事例を参考にして」では、三重県における評価システムの段階的導入と予算への活用の進化のプロセス、改革を推進した要因と論じた報告をもとに、今後の日本における可能性、特に住民、議会、首長の相互作用の重要性について議論が行われた。

第三報告の林による「公共財政管理と予算の評価」は、主に開発途上国で進められている「公共財政管理」を概観し、予算の評価を成り立たせるために必要な社会的な能力開発、特に予算システム、情報、アカウンタビリティーが三本の柱であることを強調された。

報告と質疑応答を通じ、セッションとしては「予算と評価」というテーマ設定であるが、評価制度にと どまらず、予算制度や公務員制度の見直しが不可欠であるであること、改革に向けたリーダーシップがき わめて重要であることが浮き彫りになった。

#### 共通論題セッション3「評価者倫理・スタンダード案に関するワークショップ|

三好 皓一(立命館アジア太平洋大学)

倫理規定は、専門職団体が自分たちの専門としての行動規範となるものを定めるものであり、公認会計士、弁護士、医師など、他の専門職の団体では、倫理規定あるいは倫理綱領が定められている。しかし、日本評価学会においては、評価としての倫理・行動規範は認識するも、まだこうした具体的な形での倫理規定あるいは基準規定が定められていない。評価者倫理・スタンダード策定分科会は、海外の評価学会の事例や、日本国内の他の専門職団体の事例を参照しつつ、評価活動における倫理規定の策定を目指して調査研究を続けてきた。本セッションは、「評価者倫理・スタンダード」の策定を促進するために、「評価者倫理・スタンダード」行動指針(Codes)に焦点を当て議論を行った。

セッションは、冒頭に座長より趣旨説明を行い、それを受け、一寸木会員(青山学院大学)が、「「評価者倫理・スタンダード」行動指針(案)策定に向けたいくつかの論点と考察-日本評価学会員を対象とした質問紙調査等から得られた知見より一」の演題で、「評価者倫理・スタンダード」の基本原則(Principles)案を紹介するとともに、同基本原則を踏まえて2009年、2010年に日本評価学会会員を対象に実施した質問紙調査結果及び各国評価学会における調査結果等からの知見、各国調査との国際比較、それらを基にした考察を提示し、「評価者倫理・スタンダード」の行動指針(Codes)を策定するにあたっての今後検討を要する論点を整理した。これに引き続き、岩渕会員(政策プランナー)、服部会員(青山学院大学大学院)が、「「評価者倫理・スタンダード」行動指針(Codes)案に関する基本的な考え方と今後の検討課題」の演題で、行動指針検討の枠組み、契約及び実施、レポーティング、評価結果の活用の各評価

段階の考え方、また、行動指針の策定のための今後の課題を提起した。

これらの報告を受け、討論者である西出会員(岩手県立大学)から、日本の評価では、評価を個人の行動とともに組織としての行動としてとらえることが必要であり、日本の評価における評価者について、もう一度考えていく必要があるのではないか、また、源会員(明治大学)から、個人の評価能力、チームとしての評価能力、組織としての評価能力などについても評価者としての評価倫理として議論できるのではないか、など今後の課題を踏まえた指摘がなされた。その後フロアーとの意見交換を行い、行動指針の策定の重要性を確認するともに、指針策定の基盤をより強化するためにより広い視点からの議論を続けていくことを確認した。

# 自由論題セッション報告

#### 自由論題セッション1「評価手法」

座長 山谷清志(同志社大学)

このセッションでは4つの報告が行われた。

第1報告は「地方議員の活動の評価方法開発のための基礎調査」(本田正美会員・東京大学大学院)である。2010年に各地で発生した地方議会と首長との対立・紛争から地方議会や地方議員のあり方が問われる中で、東京財団「地方議会の改革プロジェクト」(2010)を参照しつつ基礎調査を行った結果の報告である。地方議員の日常活動を細かく分析した結果、その毎日の日常における議員活動の実態がかなり詳しく描かれ、また人口規模の違いによる歳は見られないという結果になっている。これだけの知見が明らかになっただけでも有意義な報告である。地方議員は暇だという批判は陳腐化しているが、他方で本来の活動で汗をかいている議員が多いことも明らかになった。

第2報告は「評価が日本を救う」(大島章嘉会員・市民満足学会)である。適切な対象を選び、適切な指標による評価を行えば「日本を救う」ことを「幸せ度関連要因調査」で示そうとする意図で行われた調査の報告である。非常にユニークで興味深い研究であり、かつその詳細な分析と関連図表の紹介に、会場は大いに沸いた。きわめて説得力ある内容であったため、「評価」や「批判」の対象にならないのが残念である。

第3報告は「エクセレントNPO基準~課題解決としての評価~」(田中弥生会員・(独)大学評価・学位授与機構)による報告である。「市民性」「社会変革性」「組織安定性」の3つの課題を抽出し、これを基本条件としてNPOの現状と課題を論じている。そして年間500万円以下の収入規模のNPOが60%を占め(内閣府2010年)、収入と人材の不足に苦しみ続けているNPOの実態を変えるためには、いかに上記3つの課題を克服できるか、そしてその前提として各NPOの普及と「見える化」が重要であることを明らかにした。ただし大災害のあとなのでエクセレントなNPOも多い点は注目される。

最後の第4報告は「政策評価と政策分析-主に政策研究の学説史的観点から-」(宗高有吾会員・同志社大学大学院/総務省)である。政策評価と性格分析は、日本ではあまり明確に区別しないまま使用されている実態を踏まえて、両者の概念の違いを論じ、その用途がどのあたりにあるのかを実務上説明し、また

用途の違いがどこに「あるべきか」を行政学や政策学の学説から論じている。政策評価のさらなる発展をめざすためには必要な作業であろう。

筆者は司会兼コーディネーター、そして討論者を楽しく務めたが、いずれの報告も大変興味深く、いわゆる「勉強になり」、かつまた硬軟合わせて嬉しい報告ばかりであった。評価という専門性が高い学会ならではのセッションであったが、難を言えば聴衆の少なさが気になった。とてももったいない気がする。

## 自由論題セッション2「事例分析|

座長 高千穂 安長 (玉川大学)

本セッションには、21名の会員が参加し、4件の有益な報告があり、それに対して多くの質問が寄せられた。

まず、「技術協力プロジェクトの効果に及ぼすガバナンスの影響」(菊田怜子(発表)、牟田博光)では、JICAが実施している技術協力プロジェクトの効果の発現(有効性)および効果の活用と持続性(自立発展性)に対する被援助国のガバナンスの影響を構造的に明らかにすることを目的に、プロジェクトの有効性、自立発展性へ影響を与える要因の因果関係およびプロジェクトの妥当性、効率性、有効性、自立発展性に関する観測変数を明らかにし、世界銀行研究所の世界ガバナンス指標2009を用いて共分散構造分析を行った。その結果プロジェクト効果の発現という「有効性」が「自立発展性」に対して影響を与えるばかりでなく、被援助国のガバナンス状況、さらにガバナンスの影響受けた経済力、教育程度が影響を与えていることを明らかにした。より効率的に国際協力を行うための要因についての研究は今後ますます重要性を増すと考えられる。

「都道府県かつ産業毎の生産効率性と生産要素の余剰:日本の製造業に関して」(吉田素教)では、就業の中核である製造業について、生産関数の推定とDEAの実施により、都道府県かつ(24の)産業毎にその効率性と生産要素の余剰を調査・分析を目的とし、①各産業の生産関数の推定ならびにDEAに用いられるデータの精緻化、②資本設備の減価償却率が全産業を通じて一定であるという仮定の除去、③分析対象の産業がJSICの産業中分類に示された24の製造業の分析を行った。その結果、DEAについて、各地域の効率性と生産要素の余剰について新たな知見を提供した。新たな知見について、その活用の場についての研究が求められると考える。

「主要ドナー比較調査ー政策レベルのODA評価手法・体制ー」(佐々木亮(発表)、菊池美帆子、清水正治、高杉真奈、廣野良吉、長尾真文)は、外務省が実施する政策レベルのODA評価について、他ドナーにおけるODA評価の実態調査・比較を通して改善のための参考となる点を提示し、2011年より新たな体制で実施するODA評価のベースとなる「ODA評価ガイドライン」の改訂に資することを目的とし、評価体制、評価項目、評価手法、分かりやすい評価の観点から他ドナーとの比較、検討を行った。その結果、外務省としての組織決定、実務面での作業を提言に至った。このような研究は、わが国ODA評価の水準向上に直結するため、非常に重要である。世界の政治、経済、社会の環境は変化しており、研究の継続が望まれる。

「欧州における社会投資収益率(SROI)の発展とその評価」(伊藤健(発表)、玉村雅敏)は、英国を中心とした欧州におけるSROIモデルの発展と、その社会的評価を踏まえ、計量的な評価モデルがどのように非営利セクターの発展に寄与し、それを取り巻く社会の生産性の向上に資するか否かを明らかにすることを目的に、SROIの歴史、課題と解決策などについて明らかにした。その結果、日本でも成果ベースでの資金ジョセイ社会的投資への関心が高まればSROIのような計量的手法の活用が行われると予測した。今日非部会テーマであり、今後さらなる掘り下げが求められる。

#### 自由論題セッション3「教育評価/学校評価|

座長 橋本昭彦(国立教育政策研究所)

石田謙豪会員(元広島県尾道市立栗原北小学校長)「学校経営とその評価」では、同会員のこれまでの報告を受ける内容で、学校評価の改善をどのように成し遂げるかという課題について、独自の提案がなされた。学校管理者の役割、学校設置者の役割をそれぞれ7つずつ挙げ、「評価のための評価」からの脱却を訴えた。

野呂高樹会員  $((\mbox{t})$  未来工学研究所)「人材育成プログラムの評価研究 - 欧州委員会におけるMarie Curie ActionsのEx-post Impact Assessment studyを事例として一」では、EUにおける研究人材の育成と移動・確保に関する施策である「マリー・キュリー・アクションズ」に対して行われたインパクト評価の一端を興味深く紹介・分析している。息の長い計画・評価のPDCAサイクルが有効に機能することが重要なこの分野の政策評価の在り方を考えるうえで重要な研究課題であり、続報が期待される。

清水洋一会員(東京都公立小学校教員)「学校評価の活用方法-地域運営学校における学校運営委員会の役割-」は、会員の公務の都合により発表取りやめとなったため、「発表要旨集録」掲載の要旨に従って、座長から内容の紹介を試み、残余の時間を、このセッションの自由討議の時間に充てた。

佐藤由利子会員(東京工業大学)「日本と非英語圏EUへの短期留学についての評価結果の比較から」では、ともに非英語圏に属する日本からの短期留学生337人と、EU(非英語圏)からの短期留学生67人、計404人の有効回答を得たオンラインアンケートなどの結果を比較検討し、日本における高等教育の国際化の情況に対して有効かつ興味深い論議を展開した。調査手法や分析の視点が魅力的であるので、さらに手法を精緻化させた続報を待ちたい。

フロアには、毎回、初等中等教育、高等教育、学術研究の関心領域がまたがる会員が、自分の「専門外」の評価研究に関心を寄せて、有効な質疑が提出される。これは、教育という共通する分野の中で、少し異なる背景やしくみの施策・取り組みを検討し合うことで、相互比較の視点から新たな発見が共有されるという点で、学会のセッションとしては大きな実りをもたらしているもの、といえるだろう。

#### 企画委員会

企画委員会は、日本評価学会が行う各種の活動を企画、実行することをとおして、わが国の評価学の普及と発展に寄与することを目的としています。最近の活動としては以下の活動があります。

#### 1. 評価士養成講座

公共事業、学校教育、福祉、ODA等をはじめとする様々な分野の政策や事業に関するアカウンタビリティや効率化が求められており、こうした社会一般からの需要に応えるために、評価システムの導入や事業評価の実施が進められていますが、評価の専門性を持つ人材は未だ十分に育っていないのが現状です。そこで、評価に関する専門的能力を身につけた人材を養成してその能力を認定し、各分野における評価の向上に資することが講座の目的です。

評価士養成講座はすでに8回(年2回)開講され、延べ199人が受講されました。この講座の対象は、評価の専門的能力を身につけることを志向する方、および行政組織、非営利組織、教育機関、学校法人、ビジネス界など各分野で評価に携わる管理者および実務者の方々です。講座は、6つの単元から構成され、受講日数は合計6日間で、出席日数(70%以上)を満たした方には、養成講座修了証が授与されます。また「評価士」認定試験(筆記試験)に合格すると、日本評価学会『評価士認定審査委員会』から「評価士」資格認定証書が交付されます。

これまで評価士養成講座は初級レベルを想定していましたが、今年度は中級レベル、上級レベルもそれぞれ開講を開始しました。それぞれのレベルの考え方は次のとおりです。

初級レベル (評価士):評価に関する幅広い知識と技法を身につけた専門家。

中級レベル (専門評価士):各専門分野 (セクター) の評価に関する深い知識と技法を身につけた専門家。

上級レベル (上級評価士):評価に関する体系的な研究を行うことにより、評価活動の発展に寄与することができる専門家。

#### 2. 顕彰事業

日本評価学会では年一回、学会賞の選定という形で会員の顕彰を行うことにしています。学会賞の受賞者は企画委員会の下に設けられている学会賞審査委員会によって推薦し、理事会で決定されます。学会賞としては、本学会誌に掲載された原著論文等が評価学研究に大きく貢献したと認められる者に授与される「日本評価学会論文賞」、評価学研究の進歩に寄与する優れた研究をなし、将来の発展を期待される者に授与する「日本評価学会奨励賞」、評価学の発展に顕著な功績があった者に授与される「日本評価学会功績賞」の3つの賞があります。

平成23年度は10月中旬まで受賞者の推薦を受け付けています。詳細は学会ホーム・ページに記載してあります。多くの自薦、他薦を期待しています。受賞者は所定の手続きを経て決定され、11月19、20日に岩手県立大学で開催予定の第12回全国大会の際に表彰されます。

企画委員会委員長 牟田 博光

#### 国際交流委員会

2010年10月~2011年9月期の国際交流委員会の活動は、以下の通りである。

1. 各種国際会議での日本評価学会の活動紹介と仮称アジア太平洋評価協会(Asia-Pacific Evaluation Association Network=APEA-NET)設立へむけての活動

2011年2月24-25日の外務省、ベトナム計画投資省共催(日本評価学会協賛)の「第10回ODA評価ワークショップ」でのAPEA-NET設立基本合意に基づき、日本評価学会国際交流委員会APEA-NET設立検討会は、クアラルンプールにおけるマレーシア評価協会(MEA)主催の国際会議にてAPEA-NET設立趣旨および設立へのロードマップを提出することになった。4月14日マカオで開催されたアジア経営管理協会(AAMO)総会、および6月29-30日ヨルダンのアマンで開催された国際開発評価協会の研究発表大会の場においても、APEA-NETの設立の必要性、メリットを説明した。また、6月11日には、政策研究大学院で開催された日本評価学会第8回春季大会にて、従来から現在までのAPEA-NET設立に関する情報を、JES会員を含め大会参加者へ資料として配布し、今後のAPEA-NETの活動の在り方について、参加者有志と協議した。さらに、本年2月18日のUNDP本部(ニューヨーク)に引き続き、本年4月12-13日アジア太平洋経済協力会議が主催したオーストラリアのパースでの国際会議および横浜にて7月26-27日地球環境戦略研究機関が開催した国際会議と9月1-2日ソウルにて開催された北東アジアの経済環境協力会議の場においても、APEA-NET設立への参加を勧誘した。

#### 2. 全米国際問題大学院連合(APSIA)との交流

国際交流委員会は、国際協力推進協会事務所にて、2011年8月18日国際交流基金日米センター (CGP) 招請の米国国際問題大学院連合 (APSIA) 所属大学院生一行に対して、日本評価学会の活動を紹介し、特に訪問団の関心が高い、我が国の対途上国支援の現状と将来の方向、紛争後の平和構築・人間の安全保障対策、ODA評価基準について説明し、意見交換した。



#### 3. 全米評価協会(American Evaluation Association=AEA)

2011年大会への日本評価学会会員への参加・発表を、学会メールリンクおよび特設ホームページを通じて勧誘し、数名参加の予定。今後も是非、これ等海外での評価学会(協会)の研究大会へ、日本評価学会の会員の積極的参加を呼び掛けていきたい。

#### 4. 日本評価学会国際交流委員会への会員参加登録要請

2011年6月11日政策研究大学院大学で開催された日本評価学会第8回春季全国大会で、国際交流委員会への会員参加・登録を要請した。今後APEA-NETの設立と新しい国際的活動の展開のためにも、会員の皆さんの積極的な参加登録を求めたい。

#### 5. ネパール評価学会との交流

2011年2月6日、ネパール国カトマンズで開催されたネパール評価学会研究会で、ネパールと日本の評価学会の活動状況と両国における評価の現状と主要な課題について意見交換した。

6. 日本政府(外務省)とベトナム政府計画投資省共催の「第10回ODA評価ワークショップ」(2011年2月 24-25日、ハノイ市)の計画・実施に協力。

日本評価学会国際交流委員会は2001年以来、日本政府(外務省)が毎年開催してきた「ODA評価ワークショップ」の計画・実施に協力してきたが(注)、今回の第10回ワークショップにも協力した。第5セッションでは、日本評価学会のこれまで10年間の活動状況を紹介すると同時に、APEA-NET設立へ向けての段取りを説明した。本ワークショップに参加したアジア23カ国政府の評価機関代表者全員と国際機関代表者の賛同を得たことを受けて、会議終了後スリランカ、日本、ネパール、ベトナム評価学会およびサモア政府、パプアニューギニア政府代表者で構成されたAPEA-NET準備委員会を正式に立ち上げ、日本評価学会国際交流委員会の設立提案ドラフトを基に修正案を作成し、アジア太平洋諸国の現存の各国評価学会(協会)へ送付し、その意見を求め、年内のAPEA-NET設立総会の開催を期することになった。

なお、日本政府(外務省)は、2001年以来毎年「ODA評価ワークショップ」をアジア諸国の受け入れ機関(経済計画省、財務省、国家計画委員会等国によって名称は異なる)を招請し、国内外で開催してきて、今年の第10回ワークショップのベトナム政府との共催に至った。なお、このワークショップの日本との共催国は、タイ王国、マレーシア、シンガポール、ベトナムである。

国際交流委員会委員長 廣野 良吉

#### 広報委員会

#### 1. 活動方針

広報委員会は、その目的である、「評価の意義を広く普及し、学会活動の強化を支援するために、日本 評価学会および学会の行う種々の事業を広報し、学会の情報活動に寄与する」ことを目的とします。評価 学会の活動のみならず広く評価に関わる情報も会員の皆さんと共有すべくあらゆる機会を捉えてその実現 を目指します。

#### 2. 活動状況

#### (1) 「学会活動の強化と学会の健全な運営を目指して会員獲得に努めます」

会員獲得のための学会説明などのツールについては、ホームページを第一に考えていますが、国、自治体、企業などの説明ツールとしては、マス・メディアを使ったパブリシティも利用する方向で検討を進めていきます。

(2) 「広報の手段は主として学会ホームページを活用します。従って、学会からのおしらせはホームページに随時掲載します」。

引き続き、より読みやすく、タイムリーに情報をお伝えするよう検討していきます。

#### 3. 今後の方針

広報委員会の役割について見直し、進捗状況管理、新企画の検討などを行い広報面での有効性、効率性を高めていく所存です。

広報委員会委員長 高千穂 安長

# 日本評価学会第12回全国大会のご案内

第12回全国大会 実行委員会委員長 西出 順郎

来たる2011年11月19日(土) - 20日(日) に日本評価学会第12回全国大会が岩手県立 大学(岩手県盛岡市)にて開催されます。

プログラム等詳細につきましては、学会メーリングリストおよびホームページにて 随時ご案内いたします。

会員各位におかれましては奮ってご参加くださいますようよろしくお願い申し上げます。

記

1. 日時: 第12回全国大会 2011年11月19日(土)-20(日)

2. 場所: 岩手県立大学(岩手県盛岡市盛岡駅西通1丁目7番1号)

以上

#### 日本評価学会誌刊行規定

2005.2.15改訂 2002.9.18改訂 2001.9.9改訂

#### (目的および名称)

1. 日本評価学会(以下、「学会」という)は、評価に関する研究および実践的活動の成果を国内外の学界をはじめ評価に関心をもつ個人および機関に広く公表し、評価慣行の向上と普及に資することを目的として、「日本評価研究(仮名)」(英文仮名: "The Japanese Journal of Evaluation Studies"、以下、「評価研究」という)を刊行する。

#### (編集委員会)

- 2. 「評価研究」の編集は、後で定める「編集方針」にもとづいて編集委員会が行う。
- 3. 編集委員会は、学会会員20名以内をもって構成し、委員は学会理事会が選任する。編集委員の任期は 2年とし、再任を妨げないものとする。
- 4. 編集委員会は、互選により委員長1名、副委員長2名および常任編集委員若干名を選出する。
- 5. 編集委員会は、最低年1回編集委員会を開き、編集方針、編集委員会企画、その他について協議する ものとする。
- 6. 編集委員会は、その活動等について、随時理事会へ報告し、承認を受けるとともに、毎年1回学会年 次大会の場で、過去1年の活動成果と翌年の活動計画に関する報告を行う。
- 7. 委員長、副委員長および常任編集委員は、常任編集委員会を構成し、常時、編集実務に当たる。 (編集方針)
- 8. 「評価研究」は、原則として、年2回刊行する。
- 9. 「評価研究」の体裁は、B5版とし、和文又は英文とする。
- 10. 「評価研究」に掲載する原稿(以下「論文等」という)の分類は、以下の5カテゴリーからなるものとする。
  - (1) 総説
  - (2) 研究論文
  - (3) 研究ノート
  - (4) 実践・調査報告
  - (5) その他
- 11. 「評価研究」への投稿有資格者は、学会会員および常任編集委員会が投稿を依頼した者とする。学会会員による連名での投稿および学会会員を主筆者とする非会員との連名での投稿は、これを認める。編集委員による投稿はこれを認める。
- 12. 投稿原稿を上記分類のどのカテゴリーとして扱うかは、投稿者の申請等をもとに常任編集委員会が、下記の「作業指針」に従って決定する。
  - (1)「総説」は、評価の理論あるいは慣行について概観する論文とし、その掲載については編集委員会が企画・決定する。
  - (2)「研究論文」は、評価の理論構築あるいは慣行の理解について重要な学問的貢献となると認められる論文とし、その採否については次項に定める査読プロセスを経て常任編集委員会が決定する。
  - (3)「研究ノート」は、「研究論文」作成過程での理論的あるいは経験的な研究の中間的成果物に相当する論考で、その採否については次項に定める査読プロセスを経て常任編集委員会が決定する。
  - (4) 「実践・調査報告」は、評価事業の実践あるいは評価にかかわる調査の報告で、その採否について は次項に定める査読プロセスを経て常任編集委員会が決定する。

- (5)「その他」には、編集委員会が独自に企画する特集に掲載する依頼原稿や学会誌の刊行に関する 編集委員会からの学会会員への連絡等が含まれる。
- 13. 論文等は2名の査読者により査読することとし、その人選は編集委員会が行う。「研究論文」については、査読結果と編集委員会が査読者とは別に指名する担当編集委員1名の参考意見をもとに、編集委員会が掲載に関する決定を行う。「総説」、「研究ノート」、「実践・調査報告」および「その他」の論文については、査読結果にもとづき編集委員会が掲載に関する決定を行う。
- 14. 編集委員が「評価研究」に投稿した場合には、当該委員はその投稿に係わる常任編集委員会あるい は編集委員会の議事に一切参加しないものとする。
- 15. 上記いずれのカテゴリーの投稿についても、常任編集委員会による掲載の判断は可・不可の二者択一で行うこととする。但し、場合によっては編集委員会の判断で、小規模の修正による掲載も認める。「研究論文」としての掲載が適当でないと判断された場合でも、投稿者が希望すれば、常任編集委員会は「研究ノート」あるいは「実践・調査報告」としての掲載を決定できる。

(投稿要領の作成公表)

- 16. 編集委員会は、上記の編集方針にもとづき投稿要領を作成し、理事会の承認を得て、広く公表する。 (配布先)
- 17. 「評価研究」は、学会会員に無償で配布するほか、非会員に有償で提供する。 (抜刷の配付)
- 18. 「評価研究」掲載論文等の抜刷り30部を、投稿者(原著者)に無料で配布する。それ以上の部数を 希望する場合は投稿者(原著者)の自己負担とする。

(インターネット上の公開)

- 19. 「評価研究」掲載論文等は、投稿者(原著者)の了承を得て全文をインターネット上で公開する。(著作権)
- 20. 「評価研究」に掲載された論文等の著作権は各投稿者(原著者)に帰属するものとし、編集権は本学会に帰属するものとする。

(事務局)

21. 「評価研究」編集及び配布の事務は、それに関連する会計も含めて学会事務局が担当する。

(以上)

#### 『日本評価研究』投稿規定

2008.9.29改訂 2003.4.18改訂 2002.3.25改訂 2001.9.9改訂

- 1. 『日本評価研究』(The Japanese Journal of Evaluation Studies)は、評価に関する論文、論考、調査報告等を掲載する。
- 2. 『日本評価研究』は、会員間の研究成果交流の場を提供し、内外における評価研究の一層の発展に資することを主目的として発行されており、原則として会員による寄稿を掲載する。なお、依頼原稿を除き、ファーストオーサーは学会員でなければならない。また、投稿は、一時に一原稿に限るとともに、他学会誌などへ二重に投稿などのない未発表のものとする。
- 3. 投稿された原稿は、編集委員会の責任において審査を行ない、採否を決定する。審査にあたっては、 1原稿毎に2名の査読者を選定し、査読結果を参考にする。(査読者には、投稿者名を伏せて査読を依頼する。)
- 4. 原稿料は支払わない。
- 5. 『日本評価研究』に掲載された論文等は、その全文をインターネット上の本学会のホームページに掲載する。
- 6. 投稿にあたっては、投稿原稿が、①研究論文、②総説、③研究ノート、④実践・調査報告、⑤その他のうち、どのカテゴリーに入るかを明記する。ただし、カテゴリーについての最終判断は、編集委員会で行なう。「研究論文」は評価の理論構築あるいは慣行の理解について重要な学問的貢献となると認められる論文、「総説」は、評価の理論あるいは慣行について概観する論文、「研究ノート」は「研究論文」作成過程での理論的あるいは経験的な研究の中間的成果物に相当する論考、「実践・調査報告」は評価事業の実践あるいは評価にかかわる調査の報告、「その他」は編集委員会が独自に企画する特集に掲載する依頼原稿等である。

#### 7. 投稿方法

- (1) 使用言語は日本語又は英語とする。
- (2) 著者校正は原則として第一校までとする。
- (3) 英文原稿については、ネイティブスピーカーによる英文チェックを済ませ、完全な英文にして 投稿すること。
- (4) ハードコピー4部 (A4版) を提出する。その際、連絡先(住所、Tel、Fax、Email) と原稿の種類を明記すること。掲載可と判断された原稿については、必要なリライトを経た後に、最終原稿のハードコピー2部とDOS/Vフォーマットのフロッピーを用いたTEXTファイルを提出する。その際、オリジナル図表を添付すること。
- (5) 刷り上がりは最大14ページとする。これを超える場合は、その経費は著者負担とする。

- (6) 日本語原稿の最大文字数は以下のとおり。①研究論文20,000字、②総説15,000字、③研究ノート15,000字、④実践・調査報告20,000字、⑤その他適宜。それぞれ和文要旨を400字程度、英文要旨を150words程度、及び和文・英文でキーワード(5つ以内)を別に添付する。印刷は1ページ、20字×43行×2段(1,720字)とする。20,000字の原稿の場合、単純計算では英文要旨1ページを加えて合計13ページとなるが、図表の量によっては、それ以上のページ数となり得るので、注意すること。
- (7) 英文ではA4版用紙に左右マージン30mmをとり、10ポイントフォントを使用し、1ページ43行のレイアウトとする(1ページ約500words)。論文冒頭に150words程度のAbstractをつける。14ページでは、7,000words相当になるが、タイトルヘッド等を考慮して、最大語数を約6,000words(図表、注、文献込み)とする。図表の量によっては、ページ数が予想以上に増える場合もあり得るので、注意すること。

#### 8. 送付先

〒140-0002 東京都品川区東品川4-12-6 日立ソリューションズタワーB 22F

一般財団法人 国際開発センター内 日本評価学会事務局

TEL: 03-6718-5931 FAX: 03-6718-1651

E-mail: jes.info@evaluationjp.org

#### 『日本評価研究』執筆要領

2002.9.18改訂 2002.3.25改訂

- 1. 本文、図表、注記、参考文献等
  - (1) 論文等の記載は次の順序とする。

#### 日本語原稿の場合

第1ページ:表題、著者名、所属先、E-mail、和文要約(400字程度)、和文キーワード(5つ以内) 第2ページ以下:本文、謝辞あるいは付記、注記、参考文献

最終ページ:英文表題、英文著者名、英文所属先、E-mail、英文要約 (150words程度)、英文キーワード (5つ以内)

#### 英文原稿の場合

第1ページ: Title; the author's name; Affiliation; E-mail address; Abstract (150 words); Keywords (5 words)

第2ページ以下: The main text; acknowledgement; notes; references

(2) 本文の区分は以下のようにする。

#### 例 1 (日本語)

1.

(1)

(1)

(2)

(3)

#### 例 2 (英文)

1.

1.1

1.1.1 1.1.2

(3) 図表については、出所を明確にする。図表は原則として、筆者提出のものをそのまま写真製版するので、原図を明確に作成すること。写真は図として扱う。

#### 例1:日本語原稿の場合

図 ]	○○州における生徒数の推移

(注)

(出所)

表1 ○○州における事故件数	
(注) (出所)	
例 2 :英文原稿の場合 Figure 1 Number of Students in the State of 〇〇	
Figure 1 Number of Students in the State of	
Note:	
Source:	
Table 1 Number of Accidents in the State of $\bigcirc\bigcirc$	
Note:	
Source:	

(4) 本文における文献引用は、「····である(阿部 1995、p.36)。」あるいは「····である(阿部 1995)。」 のようにする。英文では、(Abe 1995, p.36) あるいは(Abe 1995)とする。

- (5) 本文における注記の付け方は、(····である¹。)とする。英文の場合は、(···.¹)とする。
- (6) 注記、参考文献は論文末に一括掲載する。

注記

1 ......

2 .....

- (7) 参考文献は、日本語文献は著者の五十音順、外国語文献は著者のアルファベット順に記し、年 代順に記載。参考文献の書き方については以下のようにする。
  - 日本語単行本:著者(発行年)『書名』、発行所
    - (例)日本太郎(1999)『これからの評価手法』、日本出版社
  - 日本語雑誌論文:著者(発行年)「題名」、『雑誌名』、巻(号):頁-頁
    - (例) 日本太郎 (1999) 「評価手法の改善に向けて」、『日本評価研究』、1(2):3-4
  - 日本語単行本中の論文:著者(発行年)「題名」、編者『書名』、発行所、頁-頁
    - (例)日本太郎(2002)「行政評価」、日本花子『評価入門』、日本出版社、16-28
  - 複数の著者による日本語文献:著者・著者(発行年)『書名』、発行所
    - (例) 日本太郎·日本花子(2002)『政策評価』、日本出版社
  - 英文単行本:著者 (発行年). 書名. 発行地:発行所.
    - (例) Rossi, P. H. (1999). *Evaluation: A Systematic Approach 6<sup>th</sup> edition*. Beverly Hills, Calif: Sage Publications.
  - 英語雑誌論文:著者(発行年).題名.雜誌名,卷(号),頁一頁.
    - (例) Rossi, P. H. (1999). Measuring social judgements. *American Journal of Evaluation*, 15(2), 35-57.
  - 英語単行本中の論文:著者(発行年). 題名. In 編者 (Eds.), 書名. 発行地:発行所, 頁 頁.
    - (例) 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 Francisco, Calif: Jossey-Bass, 177-196.
  - 2名の著者による英語文献:姓,名, and 姓,名(発行年). 書名. 発行地:発行所.
    - (例) Peters, T., and Waterman, R.(1982). *In Search of Excellence: Lessons from America's Best Run Companies*. New York: Harper & Row.
  - <u>3名以上の著者による英語文献</u>:姓,名,姓,名,and 姓,名 (発行年). *書名*. 発行地:発行所.
    - (例) Morley, E., Bryant, S. P., and Hatry, H. P. (2000). *Comparative Performance Measurement.* Washington: Urban Institute.
  - (注1) 同一著者名、同一発行年が複数ある場合は、(1999a)、(1999b) のようにa,b,cを付加して 区別する。
  - (注2) 2行にわたる場合は2行目移以降を全角1文字(英数3文字)おとしで記述する。

#### 『日本評価研究』杳読要領

日本評価学会 『日本評価研究』編集委員会 2005年9月10日決定

#### 1. 本査読要領の趣旨

本査読要領は、『日本評価研究』における掲載論文等の審査の要である査読手続きについて、投稿する会員及び査読を依頼される会員に対して解説を行い、審査手続きを効率的かつ効果的に行うことを目的として、定めるものです。

#### 2. 査読の目的と投稿者の責任

査読は、投稿原稿が『日本評価研究』に掲載される論文等としてふさわしいものであるか否かについての判定を当編集委員会が行う上で必要とされるものです。

査読に伴って見いだされた疑問や不明な事項について、必要な場合は修正意見をつけて、修正を求めることがあります。査読は、その意味で、投稿原稿の改善に資するものでもあります。ただし、修正が求められた場合においても、論文等の内容に関する責任は著者が負うべきものであり、査読者の責に帰するものではありません。

査読者は2名で、編集委員会において学会会員の中から当該分野の専門家を選び依頼されますが、学会会員以外に依頼することもあります。

#### 3. 査読の視点

査読は、以下の5つの視点によりますが、投稿原稿の種類によって、重点が違います。

- (1) テーマの重要性・有用度
- (2) 研究の独自性
- (3) 論理の構成
- (4) 実証法・方法論の妥当性
- (5) 評価理論・実践への貢献
- ・研究論文の査読については、上記の5項目全てに配慮する。
- 研究ノートの査読については、上記5項目のうち、特に(1),(2),(3),(4)の諸項目に配慮する。
- ・実践・調査報告の査読については、上記5項目のうち、特に(1),(3),(5)の諸項目に配慮する。
- ・総説の査読については、上記5項目のうち、特に(3)と(5)の諸項目に配慮する。

#### 4. 投稿に当たっての留意点

2. に掲げた査読の視点以外に、基本的な論文の完成度の問題があります。例えば、

- ●論文等として体裁が整っているか、
- 執筆要領にしたがっているか、
- ●簡潔明瞭に記述されているか、
- ●実証的なデータは適切に位置づけられているか、
- ・注や参考文献は本文と対応しているか、
- ●専門用語の使用は適切か、
- ●語句や文法的な誤りがないか、
- ●誤字脱字はないか、
- ●句読点に誤りはないか、
- ●英文要約などの英文表現は適切か、(必ずしも和文要約の直訳である必要はなく、英文としてまとまっていること)
- ●字数は規程に従っているか、

など、内容及び形式に関する留意点があります。

大学院生及び実務家の投稿において、論文としての体裁が整わないまま送付されている例があり、査 読そのものに至らないものもあるので、しかるべき指導を受けた後に投稿されるよう強く勧めます。

#### 5. 査読にあたっての判断事例

- (1) 完成度において不十分であるが掲載を考慮できる場合
  - 萌芽的な研究、発展が期待できる論文等は評価論の発展のためにできるだけ評価してください。
  - ●検証は十分とはいえないが、理論や定式化が学問の発展に有用である。
  - ●考察は十分とはいえないが、新たな理論の形成・促進に有用である。
  - ・ 文献調査は十分とはいえないが、研究の位置づけは明確である。
  - ●比較研究は十分とはいえないが、適用例としては意義がある。
  - ●考察は十分とはいえないが、社会的、または、歴史的に重要な事例の評価として意義がある。
  - ●考察は十分とはいえないが、特定の社会活動の評価として意義がある。
  - ●論文の構成や表現は適切とはいえないが、内容は評価できるものがある。
- ●論理性は十分とはいえないが、実務上の有用性がある。
- ●有意義な実践・調査報告である。
- (2) 掲載を考慮するのが困難と判断される事例
- ●問題意識や問題の設定が不明確。
- ●基本的な用語の概念の理解や分析枠組が不明確または不適切。
- ●論拠とするデータ等の信頼性が乏しい。
- ●論旨の明確さや論証の適切さがない。
- ●論文の構成、表現(用語、引用、図表等)が適切でない(または整合性がとれていない)。

#### 6. 判定

掲載についての判定は以下の4つの類型に分かれ、最終的に常任編集委員会において決定します。ただし、これらの判定は、評価できる項目や問題のある項目の多少によるものではありません。(3) 及び(4) にあるように、投稿論文の種類以外であれば、掲載を考慮できるとする場合があります。別の種類となる場合、字数の関係で、大幅に修正を要することがあります。

- (1) 掲載可とする。
- (2) 小規模の修正による掲載可とする。
- (3) 大幅な修正による掲載可とする。 但し、(総説/研究論文/研究ノート/実践・調査報告)として掲載を考慮できる。
- (4) 掲載不可とする。 但し、(総説/研究論文/研究ノート/実践・調査報告)として掲載を考慮できる。

#### 別添(1)

# **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 release evaluation studies and outputs of practical activities to domestic and international academic societies, interested individual and institutions, and contribute to the advancement and prevalence of evaluation practice.

#### 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.
- The editorial board may hold at least one meeting to discuss the editing policy, plans of editorial board, and others.
- 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.

別添(2)

# Information for Contributors (For English Papers)

Last revised on 29th September 2008

- 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 four hard copies (A4 size) of the manuscript. Contact information including mailing address, telephone number, fax number, and e-mail address, and the category of the manuscript should be clearly stated.
  - For approved manuscripts, after necessary rewriting, the author needs to submit two hard copies of the final paper as well as a text file saved on a DOS/V formatted floppy disk. 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 Hitachi Solutions Tower B 22nd Floor, 4-12-6, Higashi-Shinagawa, Shinagawa, Tokyo, 140-0002. Japan

Phone: +81-3-6718-5931, Facsimile: +81-3-6718-1651

E-mail: jes.info@evaluationjp.org

別添 (3)

# 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 photoengraved, 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 $\bigcirc\bigcirc$
	Note: Source:
	Table 1 Number of Accidents in the State of OO
	Note: Source:
	Citation of literature in the text should be, (Abe 1995, p.36) or (Abe 1995).  Note in the text should be, ( <sup>1</sup> )

(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* 6<sup>th</sup> 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.

<u>Article in Book</u>: 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.

<u>Book by two authors</u>: 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.

<u>Book by more than three authors</u>: 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.

別添(4)

#### Referee-Reading Guideline

#### 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).

#### ~投稿案内~

日本評価学会では、「日本評価研究」掲載のための投稿原稿を募集しております。1年に2回 (3月末/9月末)の投稿締切を設けておりますので、ご興味をお持ちの方は投稿規定・執筆要 領をご参照のうえ、奮ってご投稿ください。

投稿先:特定非営利活動法人日本評価学会 事務局

〒140-0002 東京都品川区東品川4-12-6 日立ソリューションズタワーB 22F

一般財団法人国際開発センター内

TEL: 03-6718-5931 FAX: 03-6718-1651

# 『日本評価研究』第11巻第2号 2011年11月10日

編集·発行 特定非営利活動法人 日本評価学会

〒140-0002 東京都品川区東品川4-12-6

日立ソリューションズタワーB 22F

一般財団法人国際開発センター内

TEL: 03-6718-5931 FAX: 03-6718-1651

印 刷 株式会社 研恒社

#### ©日本評価学会

本誌に掲載されたすべての内容は、日本評価学会の許可なく転載・複写はできません。

# Japanese Journal of Evaluation Studies

Vol. 11, No. 2, November 2011

#### **CONTENTS**

Special Issue: Benchmarking in Local Government: The Merit and the Future Prospects of Intergovernmental Comparison

Editor's Note Hiraki Tanaka, Tatsuya Ono

Future Prospects of Benchmarking in Local Government Evaluation:

The Implications of Practice in the United States Hiraki Tanaka

Comparability of the Indicators in Comparative Performance Measurement Tatsuya Ono

Municipal Evaluation Consortium:

Benchmarking for the Improvement of Local Government Service Hiroyasu Kumakura

The Administrative Evaluation of Kawanishi City and the Possibility of Benchmarking

Shigehiro Honjo

Fukui City's Performance Measurement and Municipal Benchmark System

Yoko Nakamura

The Functional Capability of Benchmarking in the Evaluation of Public Administration

-Collaboration and Evaluation Propensity in Self-interest Based Evaluators—

Junro Nishide

#### Articles

A Empirical Study of the Effects of Discussions in the Workplace on the Evaluation Process

Toru Sato

The Influence of Governance on the Effectiveness of JICA Technical Cooperation Projects

Reiko Kikuta, Hiromitsu Muta

Japan Evaluation Society