

ISSN 1346-6151

日本評価研究

Japanese Journal of Evaluation Studies Vol. 8, No. 1

March 2008

日本評価研究

Japanese Journal of
Evaluation Studies

Vol. 8, No. 1, March 2008

特集：自治体評価の10年

特集「自治体評価の10年」について	山谷 清志
自治体評価 —パイオニアの苦しみ	梅田 次郎
都道府県の評価の10年 —定量評価のこれまでとこれから—	小野 達也
都市自治体の評価：本格普及から10年後の実態	田中 啓
公共政策学・政策評価論・日本型政策評価	窪田 好男
自治体評価と外部支援指導	星野 芳昭
非営利セクターと自治体の政策評価	岩渕 公二

研究論文

市場化テストのPDCAサイクルに見る評価手法の考察 —マニフェスト評価体系からのアプローチ—	田中 弥生 工藤 泰志
自治体の評価能力向上研究 —「分かりやすい」評価説明のために—	高千穂 安長

研究ノート

評価の利用における6つのモード	吉澤 剛
-----------------	------

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
Editor-in-chief

三好 皓一(立命館アジア太平洋大学)
Koichi MIYOSHI

副委員長
Vice-Editor-in-chief

西野 桂子(ジーエルエム・インスティテュート)
Keiko NISHINO

常任編集委員
Standing Editors

牟田 博光(東京工業大学)
Hiromitsu MUTA

山谷 清志(同志社大学)
Kiyoshi YAMAYA

編集委員
Editors

青山 温子(名古屋大学)
Atsuko AOYAMA

大沢 真理(東京大学)
Mari OSAWA

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)
Tatsuya ONO

窪田 好男(神戸学院大学)
Yoshio KUBOTA

佐々木 亮(国際開発センター)
Ryo SASAKI

佐藤 真理子(筑波大学)
Mariko SATO

渋谷 和久(国土交通省九州地方整備局)
Kazuhisa SHIBUYA

鈴木 絲子(神戸女子大学)
Itoko SUZUKI

田中 弥生(大学評価・学位授与機構)
Yayoi TANAKA

松岡 俊二(早稲田大学)
Shunji MATSUOKA

村松 安子(東京女子大学)
Yasuko MURAMATSU

事務局
Office

〒140-0002 東京都品川区東品川4-12-6 日立ソフトタワーB 22F
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 薮田 みちる
TEL: 03-6718-5931, FAX: 03-6718-1651
E-mail: jes.info@idcj.or.jp

『日本評価研究』は、和文年2回、英文年1回または2回刊行します。

日本評価研究

第8巻 第1号 2008年3月

目次

特集：自治体評価の10年

山谷 清志

特集「自治体評価の10年」について……………1

梅田 次郎

自治体評価 ―パイオニアの苦しみ……………3

小野 達也

都道府県の評価の10年 ―定量評価のこれまでとこれから― ……19

田中 啓

都市自治体の評価：本格普及から10年後の実態 ……39

窪田 好男

公共政策学・政策評価論・日本型政策評価 ……59

星野 芳昭

自治体評価と外部支援指導 ……73

岩渕 公二

非営利セクターと自治体の政策評価 ……83

研究論文

田中 弥生 工藤 泰志

市場化テストのPDCAサイクルに見る評価手法の考察
―マニフェスト評価体系からのアプローチ― ……95

高千穂 安長

自治体の評価能力向上研究 ―「分かりやすい」評価説明のために― ……115

研究ノート

吉澤 剛

評価の利用における6つのモード ……125

第8回全国大会開催報告

第8回全国大会プログラム（実績）	141
共通論題セッション報告	145
シンポジウム報告	150
自由論題セッション報告	151

委員会活動

企画委員会	156
研修委員会	157
広報委員会	158

春季第5回全国大会のご案内	159
---------------	-----

特定非営利活動法人日本評価学会定款	160
日本評価学会誌刊行規定	170
日本評価研究投稿規定	172
日本評価研究執筆要領	174
日本評価研究査読要領	177

【巻頭言】

特集「自治体評価の10年」について

山谷 清志
同志社大学

地方自治体で政策評価、行政評価（以下「自治体評価」）が試みられてから既に10年以上の年月が過ぎている。そして、この自治体評価への取り組みには積極的に取り組んだ自治体と消極的な自治体があったのは当然であり、どの自治体が積極的であったのかは、他の改革にも熱心であった「改革派首長」の名前とともに既によく知られている。もっとも、積極・消極、熱心・不熱心の判断は観察者の視点に左右され、首長のパフォーマンスによっても影響される。学術研究としては、単なる印象論にもとづく判断ではなく、客観的な分析が必要であろう。そこで4つの類型論からはじめたい。

第1の類型は1996年頃に導入を検討し、1997年頃から制度設計をはじめていた自治体で、評価に積極的な期待をしたところである。代表が三重県庁で、「評価を梃子に県庁を変え、県庁が変われば地域が変わる」というようなスローガンで評価に期待を寄せていた（「さわやか運動」に象徴される図1の状況である）。第2の類型は1998年頃から自治体改革に積極的に取り組み、多くの改革手段を取り入れ、そのさまざまな改革の一つとして評価も導入したところである（図2）。岩手県庁や宮城県庁、北海道庁、秋田県庁などがよく知られている。前後して取り組んだ改革、たとえば行政経営品質向上、公共事業再評価、独立行政法人や市場化テスト、マニフェストのように、これらの改革のうちのかなりの部分が、自治体評価と何らかの関わりを持つてくる。

図1 三重県庁型

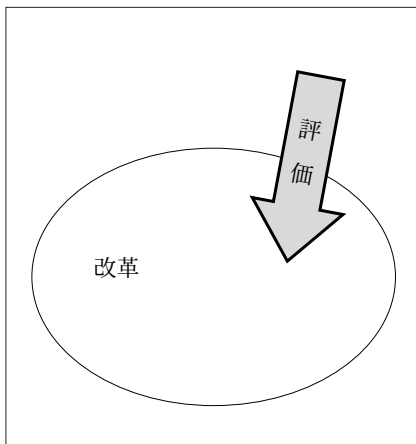
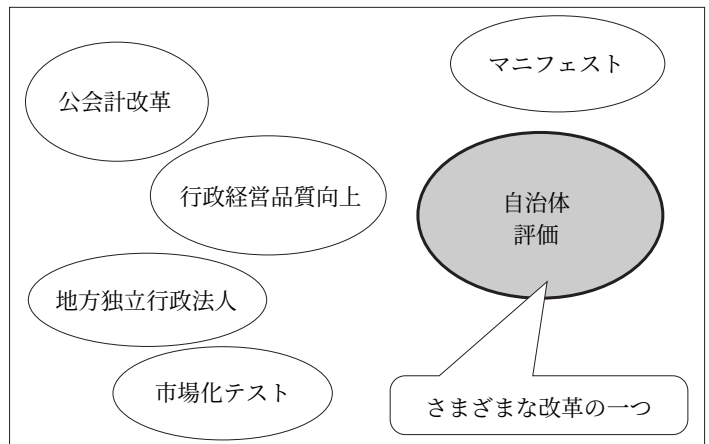


図2 岩手県その他の自治体型



これら積極派の2類型に分類された自治体は、評価制度導入後にさらなる改革を関連して行なう。たとえば外部の有識者や住民の代表による委員会の設置、事務事業評価制度を発展させた施策評価制度のチャレンジ、青少年・助成・健康・教育などの専門評価の構想、NPOとの連携、地域住民の参画とエンパワメントなど、評価に連携させて新しい試み、しかし冷静に考えると「イバラの道」に取り組みはじめる。小手先の改善にとどまらない「永続革命」戦略を採用した「原理主義者」になったかのような錯覚を一般

職員に抱かせたのである。もちろんその次に来るのは「評価疲れ」であった。

他方これらとは違い、積極的でない自治体もあった。それが第3の類型であり、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「導入すればそれでよい」と考えていた自治体である。つまり導入することが目的であり、良くて「行政による、行政のための評価」、悪くすると「強制された苦役」に墮落する。そうした自治体が評価を導入した理由は、先に導入した同規模自治体に対する横並び意識、評価に素人の首長の指示があった。とにかく導入しようとまじめな職員が数人がかりで取り組み、評価の先進自治体を参考に、コンサルタントのアドバイスを受け、一応それなりの形の評価システムを構築し、器用にまわしている。ただし、公務員として妙に生真面目なあまり手続や書式の厳格さにこだわり、「評価マインド」の良きいい加減さ、すなわち「試す」けれども「無意味なことはやらない」、縦割り行政に「拘泥しない」、部局をこえ住民をまじえて「みんなでオープンに議論する」意識は希薄である。「昨日と同じ明日が続く」と思いたいのである。したがって、評価の「面白さ」はなく、形骸化しマンネリ感に沈むことになる。

第4番目の類型は、まったくその気がなく、形だけは事務事業評価を入れたが、そのほかは何もしなかった自治体である。基礎自治体にこの類型がなぜ多く発生したか、その発生原因は二つ考えられる。ひとつは「評価どころではない」自治体である。「三位一体の改革」と市町村合併の激動に翻弄され、夕張市の財政破綻にショックを受け、地方自治体財政健全化法（2007年6月成立）による公営企業・第三セクター会計との連結決算に驚愕し、慌てて市場化テストや指定管理者制度に飛びつき大ヤケドした自治体である。もう一つはもっと悲惨な基礎自治体であり、ここでは公金流用・違法起債・特殊勤務手当二重払い・道路特定財源（使い切れないために）不正流用・徴収した市税の行方不明・市職員の市営住宅家賃と市営保育園料滞納・ヤミ退職金とヤミ年金・入札妨害職員・巨額の残業手当不払い・国庫補助金のずさんな申請受給・ハコモノ建設への異常な固執などの「症状」が顕著である。この第4番目の類型に共通しているのは‘bad-governance’か、前近代行政の‘mal-administration’の残滓で、評価以前の段階にある。

都道府県と市町村を10年余観察して発見した4類型の先に見えてきたのは、卑近なたとえで恐縮だが、3種類の自治体の存在である。最初は「言われなくてもちゃんとやる」自治体、つまり第1と第2の類型で、ここはまず大丈夫である。ただ大丈夫さを保つには自ら率先してイノベーションを繰り返す必要があり、脳力と体力が要る。二つめの種類、第3の類型は「言われたことはちゃんとやる」自治体である。もっとも「言われないこと」はやらないので、評価の形骸化と沈滞が待っている（‘capacity-building’と人事の巡り合わせ次第で良くなる可能性はある）。三つめの第4の類型には評価は要らない。「言われたことすらちゃんとできないくせに、変なことをする」からである。合法性やコンプライアンスがきちんと担保されないのに、効率や有効性を主張するのは無意味である。自治（self-government）能力も疑わしい。

評価を通じて見えるのはガバナンス能力の有無である。あるいは自治体職員の「志」の高低である。そして、自治体評価は政策や活動を評価するツールであることはもちろんだが、実はその自治体が今どんな状況にあるのかを判断するメルクマール、リトマス試験紙の機能も持っている。『日本評価研究』8巻1号の特集のタイトル「自治体評価の10年」は、平凡で月並みな印象を受けるが、実は単なる評価技術論にとどまらない今後の研究への示唆を含んでいると自負している。

【研究論文：依頼原稿】

自治体評価 ―パイオニアの苦しみ

梅田 次郎

JMAC構造改革推進セクター

jirou_umeda@jmac.co.jp

要 約

三重県が1995年から取り組んだ行政改革は、事務事業評価システムをその根幹に据えるという前例のないものであった。このシステムが自治体における評価の先駆となった。筆者は当初からその実務上の責任者として携わってきた。

本稿では、その誕生に至った要因と初期の導入過程を県庁の内部からたどりながら、パイオニアとしての苦しみについて実務体験を中心にまとめた。官僚組織に内在する組織的抵抗とそれを破ろうとする職員の変革力とのせめぎ合いがあったのである。官僚組織の病理現象の改善なくして、評価の実効は期待できないと考えている。

キーワード

地方分権改革、意識改革、政策形成、業績測定、官僚組織

1. 本稿の目的

1995年から三重県において取り組んだ事務事業評価システムの担当責任者であった筆者は、これまでにそのシステム誕生の経緯、考え方と進め方、発展の過程、運用展開による県庁内外における変革の状況などについて考えをまとめてきた（梅田・竹内1999、梅田2000a、2000b、2001、2002a、2002b、2004、梅田・西城2003）。

本稿は、あらためて第1になぜ三重県で1995年という時期に取り組み始められたのか、その経緯をたどる。第2にその導入期の当初3年間におけるいわゆるパイオニアの苦しみについて筆者のこだわりも含めて実務体験を中心にまとめるものである。

2. なぜ三重県で1995年という時期に取り組み始められたのか

(1) 地方分権推進の流れ

1990年代初頭日本のバブル経済崩壊後の1993年6月に衆議院及び参議院で「地方分権の推進に関する決議」がなされた。その決議書には「地方分権を積極的に推進するための法制定をはじめ、抜本的な施策を総力をあげて断行していくべきである」とあり、いよいよ地方分権への取り組みが本格化する流れが決定的になった。

政府は1994年2月「今後における行政改革の推進方策について」を閣議決定し、規制緩和と地方

分権の推進を柱として進めることとした。これを受けて同年10月7日自治事務次官通知「地方公共団体における行政改革推進のための指針の策定について」が各都道府県知事および各指定都市市長に対して出された。その中で行政改革推進本部の設置や1年以内を目途とした行政改革大綱の策定等が求められていた。具体的な項目としては、事務事業の見直し、時代に即応した組織機構の見直し、定員管理、給与の適正化などそれまでと変わらないものであったが、今回は地方分権の推進という国と地方の関係を問い直す要素が含まれていただけに、三重県としても明確な取り組み姿勢を見せた。ちょうど1994年7月から三重県伊勢市で開催していた世界祝祭博覧会（入場者351万人）の11月終了をとらえて、1995年1月に人事異動が行われた。筆者は博覧会事務局事業部長として、パビリオン、イベント、営業の担当部長であったが、この異動で総務部参事（4月からは審議監）となり、その行政改革を担当することとなった。

筆者は1970年代末人事課に在職して行政改革を担当し、当時流行していた事務事業総点検の作業を体験していた。各課より大量の資料を提出させたが、集まってきた資料を見ると「現状維持」ばかりで、途中でチェックをやめてしまい、あとは組織を削るなどの改革に終わったことがあった。この苦い経験への反省が、後で生まれる事務事業評価システムへの取り組み姿勢に生かされた。

筆者にとっては10数年の事業畑を経た後の久しぶりの行政改革担当で、少し勉強しなければと思い本屋へ行った。そこで、ちょうどその1月に翻訳出版されていた『行政革命』（Osborne and Gaebler 1992）を入手し、異動先の人事課の部屋の隅で、このような行革がしたいものだと興味をそそられながら読んでいた。人事課は県庁の中で最も保守的な雰囲気をもっているもので、人が近づくことと引き出しにしようという有様であった。この本が後でいろいろな動きを生み、形をつくる重要な契機のひとつになった。

地方分権推進法案は、1995年2月に閣議決定され、国会に提出、同年5月に可決成立、7月に施行された。いよいよ自治体職員の中央官庁依存意識からの脱却と政策形成能力の向上が強く叫ばれるようになった。

(2) 北川知事の就任

1995年4月9日統一地方選挙が行われ、北川正恭氏が、それまで副知事であり前知事路線継承の対立候補尾崎氏を1万3千票弱の僅差で破り当選した¹。北川氏は三重県議会議員を1972年から11年間、衆議院議員を1983年から11年間を経ての就任であった²。

4月21日の知事就任から1週間後知事との最初の行政改革の打ち合わせが行われた。選挙公約からすると改革の結果を早く出さないといけないのではないかとの憶測から、秋ごろには改革結果を出す、いわば速成の改革プランを提示した。これを見るなり「こんなものでは駄目だ」と即座に否定された。それはわれわれに対する怒りに近かった。そこで間をおかずに2回目の打ち合わせを行い、もう少し時間をかけ1年がかりで取り組む案を提案した。知事からその改革案の個々の文章について「これは何をやるのか」と細かく問われたが具体的に答えられないことが多く、役所言葉が踊っていることを知らされた。その2回目の打ち合わせのときに、知事の考えている改革の方向が示された。

- ①情報革命、経済のグローバル化の進展は、公平公正なルールと透明性の高い組織運営を求めている。行政組織も、官の論理、権力を持つものの論理から、「生活者」の論理へ転換しなければならない。
- ②長く続いた中央集権システムの中で思考停止とでもいうべき状態になっている職員の「意識改革」がすべてのスタートである。
- ③いわば独占企業である県庁の組織文化を抜本的に見直し、もうひとつの県庁と競争しなければならないような状態におきたい。
- ④そのような改革を推進するために、民間の経営手法を導入したい。特にその事業管理手法を導入したい。
- ⑤そのためには民間経営コンサルタントを導入したい。役人の作業だけでは駄目だ。

生活者という言葉キーワードにすること、職員の意識改革にまず取り組むこと、民間経営手法を導入するため民間経営コンサルタントを入れることが最初の指示事項であった。

「行政改革に民間経営コンサルタントをいれる」という発想は筆者たちにはなかった。部分的に調査作業等を依頼することはあるが、改革の作業は職員で行うものというのが常識であった。それまでにそういうコンサルタントとの接触もなくどうしたものかと頭の隅で考えていたとき、知事があるメモをテーブルに置いた。こういう行政改革が面白そうだとのことだった。見るとそれは『行政革命』の目次であった。筆者はすでに読んでいて興味を持っている旨を告げた。偶然の出会いであった。打ち合わせ終了後、すぐさまその翻訳出版先に電話を入れ、具体的な案作りを進めた。

このようにして星野芳昭氏（現在、JMAC構造改革推進セクター事業責任者）との接触が始まり、5月連休前にFAXが送られてきた。星野氏からは「政策・事業の評価システム」という案の提示があった。私は、「これだ！これまでの行政改革で出来なかったことをやろうとするなら、これを根幹にすえてやることしかない。これまでの伝統的な手法を捨てることだ。予算が減る、増えるは、その結果にしか過ぎないというシナリオだ。」と直感した。そして、システムの名称については、役所ではすべてのことを表す言葉として「事務事業」が定着しているので、これを「事務事業の評価システム」に改めた³。

筆者はかつて事務事業総点検で失敗体験をしているので、今回はその二の舞になってはいけなさと感じた。事務事業総点検のときは、価値観の転換を促すわけではなく、ただ単に点検様式を決めて通知し、期日までに提出させるという安易なやり方であった。今回は、生活者を起点に事務事業を見直すことを明示し、職員の意識改革を呼びかける。そのためにはこれまでの役所の組織文化、官僚組織の病理現象にまで踏み込んでそれを直すやり方、すなわち大々的な評価作業しかない。3回目の打ち合わせで説明した。打ち合わせ2時間で知事は「これでいこう！」と即決。5月中旬であった。午後7時ごろ打ち合わせを終えた筆者と担当課長ら3人は県庁のロビーで今後の進め方について話し合っていた。やるべきと考えてはいたがあっさり決まってしまったことに快感と戸惑いがあった。今後の対応についてとりとめもなく話し合っていたがなんと1時間半の立ち話をしてい

たのであった。

(3) さわやか運動の展開

知事の決定を受けて、「生活者起点の行政運営」を基本目標とし、地方分権時代にふさわしい県政の確立を目指した具体的な展開を①目的・成果志向での事務事業の見直し体制の確立、②政策に反映される事務事業の評価システムの確立、③先進県としての行政目標挑戦体制の確立の3段階の展開イメージを構築した（三重県 1995）。

筆者は、こういう方向を指し示す言葉としていわゆる「行革」という言葉を使いたくなかった。行革＝削減であり、この固定観念からは、時代にあった創造的な変革イメージは出てこないと思った。県庁の組織文化、職員の行動様式を根本的に変えようとする質的な改革は一朝一夕に実現できるものではなく、また強引に権力的に進めてみたところで身に付かず、すぐ元に戻ってしまうものであるから、職員の意識改革を中心に据えた運動論として展開することが望ましいと考えた。そこで、「何々運動」としたいと考え、担当課の職員に呼び名を募集した。応募案のなかに「さわやか運動」というのがあった。後日、さ＝サービス（＝行政の価値を高める）、わ＝わかりやすさ（＝生活者を起点に行政を見つめる）、や＝やる気（＝一人ひとりが目標を立て、挑戦する）、か＝改革（＝既成概念を捨て、白紙で考える）と整理されたものである。評価システムといい、さわやか運動といい、案を見たときに、筆者は「これだ、これしかない」と思った。が、同時に「正直この訳の分からないものに突っ込んでいっていいのだろうか。普通の役人の流儀で行くなら、避けるという判断だろうな」と不安ただならぬものがあったことも事実である。「行革」という言葉を前面に出さない「さわやか運動」は知事の英断で採用された。知事は就任後初めての1995年6月県議会の所信表明で、55年体制の崩壊という大きな歴史の変革期を迎えている現在、旧来の行政改革とは全く異なった視点で取り組むことが必要であるとして「さわやか運動」を提唱し、その目標運動期間を3年間と明示するとともに、事務事業評価システムの確立をその根幹に据えた。

推進体制として1995年7月20日には、総務部とは別の知事直属の地方分権・行政改革総括推進室が設置され、筆者は引き続きその実務上の責任者となった⁴。

以上が事務事業評価システム（以下「評価システム」と略す。）の誕生の経過である。

3. パイオニアの苦しみ

あらためて初期3年間（評価表の公表まで）を振り返り、どのように推進していけばよいのか思い悩んだことや、どのような誤解や抵抗感と戦ってきたかなどについて、あらためてまとめ直しをしてみたい。そこには挑戦する役所の悩み、すなわち官僚組織に内在する組織の抵抗とそれを突き破ろうとする職員の自発的な変革力とのせめぎ合いのドラマがあったのである。

(1) 外部支援に予算をどれだけ使うか

予算を計上するとき、筆者にはかなりお金をかけないとコンサルタント側も本気にならないだろうという予測があった。当時総合計画作成委託費は相場が2千万円から3千万円であったので、この2倍は計上する必要があると考えた。6千万円を超える予算要求案は財政課を通過し、知事の決裁もおりた。この金額の大きさがそれ以降のコンサルタントと役所の人間との付き合い方にまったく新しいスタイルを生んだ。

役所の事務部門で「コンサルを使う」といえば、単なる作業委託のほかにはプロジェクト基本構想の策定や意識調査など「一定の業務をそつなく代行してもらって、きれいにまとめて、印刷してもらう」というケースが圧倒的に多いのが実態であった。そこで、「評価でコンサルタントの支援を受ける」と言えば、ほとんどの人が「評価を代行してもらうこと」、あるいは「コンサルタントに評価されること」と誤解してしまった。しかし、今回のコンサルタント導入は、これまでの役所の常識的なコンサルタント導入イメージとは全く違い、システムの構築、導入支援だけでなく職員の意識を変えることを究極のねらいとした。結局3

年間で2億円以上の経費をかけ、3,000人を超える職員（部長等の幹部を含む。）に対して、何回も研修・説明会や点検会を実施した。「職員とコンサルタントとの対話」にお金と時間をかけたのである。今から振り返ってみると、結果的にはこれがすべての改革の基盤となったが、職員の抵抗感はかなり強く、理屈で説明しても分かってもらえず、実践を通じ時間をかけていくしかなかった。

また、筆者はコンサルタント側が従来型のイメージで取り組んでもらっては困ると思い、始める前に次のように言った。「よくある、分厚い報告書は要りません。今回は、職員の意識が実際に変わることが実績報告です。」

筆者は、役人が役人をたたくことの限界を強く感じていたので、コンサルタントという外部者との議論に大いに期待をかけていたが、実は半信半疑であった。「初めてのこの試みは果たしてうまくいくのだろうか。経費も相当かけているので、改革への効果が出なければ壮大な規模でまったく無駄な作業をしたことになってしまう。」そのプレッシャーは日に日に増大していった。

「何を言っているんだ。行政のこと、何も知らないくせに！」筆者がドアを開けて点検会の様子を見ようと入った途端に聞こえてきたせりふであった。そこには、評価表を前にして激しく議論している光景があった。プロとしてのコンサルタントと、プロとしての土木事務所長の間に激しい火花が飛び散っていた。私は、なぜか、すぐ部屋を出たが、「これだ！ これでいいのだ！ これが改革の始まりだ！」と心の中で叫んでいた。

激しく言い争う職員や黙りこんでしまっている職員がいるかと思うと、なるほどと目を見開かされて談笑している職員もいた。さまざまな反応と場面が毎日繰り返されているのを見て、筆者は職員とコンサルタントが対峙する構図を次のように理解することが出来た。

評価表を前にしてAコンサルタントは聞き手で、B土木事務所長は話し手である。Bが話し続けるうちに、聞き手役のAは次第にBを正確に映す鏡になっていく。話し手BはAという鏡に映った自分の姿を、自分の姿とは認めない。それまで自分の姿を映していた鏡は、もっと自分をきれいに見せてくれる鏡、心地よい鏡であったからであ

る。Aは聞き手に徹することによって、「これがあなたの姿なのです」と示し続ける。鏡に映る姿を自分のものと認めたくないBは、耐えられず、Aに対して怒り出す。「行政の事を何も知らないくせに！」と言いたくなるのである。

行政の内部事情のことを知らないからこそ、正確な鏡になれる。内部事情を知りすぎていると、鏡がそれを踏まえる分だけ歪む。歪んだ鏡に映しているからこそ、これまで心地よくきれいに見えていたにすぎなかったのではないか。自分の声を録音・再生すると、こんな声だったのかと愕然とするのに似ている。

コンサルタントは、職員との対話の場で評価の代行をするのではなく、評価作業をする人の鏡になってサポートする。コンサルタントを医者にと考えると、患者の自己治癒力を高めるように仕向けていく医者、患者が自らの病理に気づき、自らが改革案を書き、それに向けて行動し始めることを上手く促していく医者といえた。

(2) 民間経営手法の援用が行政現場に受け入れられ、有効に働くのだろうか

1995年7月地方分権推進法の施行に伴い地方分権推進委員会が発足し地方分権に向けて具体的な検討が始まったころ、同月27日にさわやか運動のキックオフ大会を開催した。その大会の司会をしていた筆者は、知事のいる前で「皆さん、知事が要らなくなるくらいに評価システムをしっかりとやっていきましょう」と呼びかけた。皆どっと笑ったが、言い放った筆者にこの評価システムが方法論としてどこまで有効か、また職員に果たして受け入れられるものかどうかについての確信はなかった。

8月から取り組んだ幹部研修の中で、その可能性を探ることから始まった。幹部研修は部長・次長級研修70名、課長級研修530名で、1泊2日の研修を2回行った。これらの研修は、翌年2月まで行われたが、評価システムの可能性や実効性について、一挙に確実な手応えをつかんだわけではなかった。いわば手探りで、恐る恐る方法論を提示しながら第1段階として可能性を予感し、第2段階として実効性について見通しがつく段階に至ったの

であった。

評価システムの可能性を予感させたのは、部長・次長級研修における、三重県の行政組織・運営上の問題点についての受講者の共通した現状認識であった。次のような意見が出された。

- ①政策理念が不明確で、総合検討や全庁的議論が不足している。
- ②議論を避け、事なかれ主義になっている。
- ③本庁・財政当局主導の政策決定、予算重視、決算軽視となっている。
- ④行政の成果に関する評価システムが欠如している。数値管理が不十分であり、言葉で逃げている。
- ⑤業績評価制度がない。減点主義による人事管理で職場は沈滞気味である。

以上の現状認識は、一般的に官僚組織の病理現象といわれるものであるが、私は、その内容の率直さとともに、全てのグループが同じようなまとめ方をしていることに驚きを禁じ得なかった。と同時に、これだけの現状認識があれば改革に踏み込めるかも、という希望的な感触を初めて持った。

さらに、今後の改革の方向として次のような指摘が出された。

- ・政策の計画的推進、財源の重点的配分のための「効果の評価」の実施
- ・各部・各課所の自主性と責任体制の確立
- ・職員の意識改革（誰のための県庁かの確認、使命感、プロ意識、コスト意識）
- ・目標管理、プロジェクト管理、評価システムの確立
- ・業績主義の人事管理システムの再構築

このような改革の方向は、行政改革が叫ばれる度に議論されてきたもので、別に目新しいものではなく、当たり前のものである。しかし、これらの改革を本気で実現しようとした実例がどれだけあったかは、疑わしい。行政組織は、官僚制の原理に基づいて機能的に運用されるべきであるが、その官僚制そのものを徹底することからくる病理現象に対する治癒力を開発しなければ、これまでの行政改革の限界をこえることはできない。従来の量的削減を中心とする行政改革の手法に幾度となく空しい思いを抱き続けてきた筆者にとって、上記のグループ討議で出された課題認識と改革の

方向は、評価システムが官僚組織の病理現象に対して個別・部分的でなく、システムとして働きかけ、官僚制の合理性に対抗して内部から民主化する機能を持つ可能性を予感させるものであった。

さらに、評価システムの作業が実務として進めることができるとの手応えを感じたのは、課長級研修における評価システムの試行状況及び体験者の次のような感想、意見からであった。

- ①行政運営に使命・目的・成果の評価が重要であることを学んだ。事務事業を行政使命から成果指標まで関連付けてみる手法を勉強した。特に成果志向で業務を遂行するのを感じた。
- ②行政としてやるべきことは何かについて考える意識ができた。基本、原点に戻れた。自分の部門の問題点・欠点を再認識した。政策形成や部門改革についてまだまだ不十分であると理解できた。
- ③県行政の新たな取り組み手法だ。現実的に難しい面もあるが、単なる研修でなく、実際に取り組んでほしい。

アンケート用紙に書かれたこういう意見に元気づけられながら推進できたのである。

しかし、評価システムの意義を認めながらも運用上の問題点、疑問を提示する意見や、結論としては導入に反対するという意見も多く出た。会議の席上、反対意見が出ると、誰も反論しようとはせず、それが決定的多数意見であるかのような雰囲気その場に満ちた。

- ①すべての業務を画一的に評価することに難しさがある。実質的な効果は疑問である。
- ②効果測定のための要因が複雑すぎて数値化は難しい。成果指標を客観的に定めることはきわめて困難である。
- ③現実には当てはめるとき、既成のシステムを変えることがいかに大変か、これに費やすエネルギーをどこから生み出すのか自信がない。研修としてならよい。

導入に消極的な意見は、こんなことをしても所詮現実是不変、変えたくない、ならば膨大な事務量を考えれば導入したくないというところに集約された。「所詮現実是不変」という意識は、いわゆる55年体制の下、政官財の護送船

団方式による政策展開を支えた長年の中央集権システムの中で「明日の補助金、明後日の許認可」を求めて毎日残業しながら懸命に働いてきた自治体の職員の総括といえた。仕事に精通し、現在の事業の経緯をよく知っている職員ほど、この評価システムは無駄な作業に見えてくる。こういうものは、研修の次元で対応しておくことが生活の知恵であるとするのであった。

評価システムによって事業の不合理性が明らかになることが予想されればされるほど、生じるであろう混乱、痛みに敏感にならざるを得ないが、「だから、改革に飛び込まない」ではなく、「だから、飛び込もう」と呼びかけ続けた。

この一連の研修で、数は少ないが欠席者があった。それを知事に言い付けたところで逆に皆から反発を買うかもしれないと思い黙っていたが、その年の12月、知事はどこかで聞きつけたらしく「部長研修で欠席した者がいたのではないか」とお尋ねがあった。「部長級で3名」筆者はもちろんその旨を告げた。さらに「ほかにはいなかったのか」との問いに「課長級は、管理部門の課長2人です」と筆者は答えた。知事からの指示は、その人たちだけを対象に同じ研修を行えというものであった。このニュースは県庁中に広まるところとなり、以後欠席者はいなくなった。

(3) 組合の反対

1995年10月ごろ三重県職員労働組合から電話があり、評価システムについて話したいということであった。評価システム構想を発表し、さわやか運動に取り組み始めて4ヶ月がたった。

役員15名ほどが会議室に集まった。組合からは、「これはどうも予算削減の道具ではないか」というもっともな疑念が強く表明された。予想通りであった。筆者は黒板に縦と横の線をひきながら話し始めた。

「まず縦軸として、行政の「対象」を切り分けてください。たとえばこれまで陽の当たらなかった対象を明確に掲げるのです。次に横軸として行政の「意図」を切り分けてください。従来型の意図ではなく成果志向の新しい意図も掲げます。すると新しいマトリックスができますね。新しい対

象と新しい意図のクロスするところには新しい成果志向の政策が生まれるでしょう。また、そのマトリックスの表に既存の政策を埋め込んでみてください。そうすると必ず空欄ができますね。その空欄のところへ政策を立案するのです。こうして新しく政策形成した後、既存の政策と比較します。人と金には限りがあるわけですから、あれか、これかの選択をすることとなります。結果、どちらかがスクラップされるということになりますが、これは従来の行革で言うスクラップとは質的に違うのです。これまでは「スクラップ・アンド・ビルド」でしたが、これは「ビルド・アンド・スクラップ」ともいうべきものです。評価システムはその「ビルド・アンド・スクラップ」のところで構築しています。このシステムを使って政策形成し予算要求してほしいのです。成果の数値目標を掲げて予算要求するのです。これまではこういう風に要求しなかったから、予算もつかずに皆さんは愚痴ばかり言っていたのではないのでしょうか。それを認めさせ、予算をつかせて、政策を実行するのです。実行後は目標数値を達成したかどうか検証して、次年度の改善につなげていくのです。評価システムはただ単に予算を削る道具ではありません。この道具であなたの主張、政策理念を実現させてほしいのです！」チョークを持つ指に力が入り、かなり大きな声で呼びかけていた。この立ち上げの時期に、組合に反対されたら評価システムの作業が困難になると直感していたからである。

議論が尽き始めたころ、一番後ろの席に座っていた女性役員が手を上げた。私は反射的に腕時計を見た。2時間経過していた。次の瞬間私の頭の中にひとつのシナリオがひらめいた。「これは意識改革のための研修という意味もありますので…」と竜頭蛇尾に終わらせるシナリオであった。「やはり、これは削減の道具だから反対する」という発言を予想しての反応であったが、その直後に聞いた言葉は、なんと「私たちは、応援します」であったのである。

その言葉があつてその後のシステム構築の作業が持続できたとつくづく思う。それを聞いた瞬間の風景をいまでも感動的に思い出す。

さらに筆者は叫び続けた。「地方分権の時代だ、

政策形成だと叫ぶその口で、同時にスクラップばかり叫んでいないか。目指す分権社会の自己決定、自己責任の政策づくりをするには、性善説に立つべきではないか。現場が一番良く知っているのだから、現場が政策を形成するようにもっていかなければ意味がない。現場に任すという意味で、まずは役所の中では職員参加をめざす。そしてひいては、地域に任す住民参加が基本になる。そのためには、集中管理に巣食う性悪説ではなく、性善説に立つことが必要です。性善説にたって運用していけば、評価は政策形成を現場に任せていくための道具となります。任すときには成果目標つきで任す。その数値目標に照らして達成度を測定して評価を行うのです。」

この会議以降、筆者らはすべての情報を組合側に渡すこととした。味方になってもらおうとしたのである。組合も取り立てて反対しなかった。

(4) 本格導入への突入と職員の抵抗

評価システムを本格導入するかどうかは1996年度予算編成時に決まる。その直前の1996年1月4日には恒例の知事訓示が講堂で行われたが、それは、筆者の予想をはるかに超えたものであった。すなわち話のはじめからおわりまで、評価システムの話に終始したのである。こういうことはかつてなかったことだった。終了後、幹部職員らは筆者を見つけては「原稿書いたんだろう」と揶揄したが、一切そのようなことはなかった。知事の断固たる決意表明に筆者自身驚き、期待と不安の相克する心理状態になっていた。

その後予算も正式に決定され、1996年4月から評価システムの本格的な導入に入っていた。

ちょうど同年3月29日地方分権推進委員会から中間報告が発表され、国と地方の関係を「上下・主従」から「対等・協力」の関係に改める方向性が明確に打ち出された。評価システムの狙いはそのための意識改革と政策形成に向けてであることを強調した。

導入方式は、県庁の組織文化に切り込むために部分的・モデル的導入ではなく全庁一斉方式とした。1年目は本庁全体を対象とし、2年目は出先機関にも拡大した。作業にかかわる職員数は、1年

目が600名、2年目が1,700名であり、参加者の出欠調整事務だけでも大変であった。4月早々から手引書による研修・説明会を行い、その後各課で書き込みをしてもらったシートについてコンサルタントと担当課との対話による点検会をおこなっていった（三重県 1996b）。点検会のような職員の理解を促すサポート体制をとることは煩雑で手間がかかるが、これなしに権力的に進めるだけではすぐさま導入は頓挫してしまうと考えていた。

職員の反応としては、評価システムの意義・効果に理解を示しつつも、作業に直接携わる職員の立場から業務量の増大、成果指標の数値化の困難性、評価システムの実効性等に関する意見や悩みが続出した。

新しいことをやろうとすると、完璧な100点満点を追い求める議論をする職員が多数出てくる。説明会場はほとんどそういう人たちでいっぱいになったが、じっと議論を聞いていると、「やりたくないから、そう言っているのではないか」と思うことが多かった。役所の無謬性神話の中で育ってきた職員間の議論では、100点満点でない新提案は時期尚早となり、結局廃案となるのが普通の道筋であった。改革しようとする人たちは、たとえ80点の手法でも藁にもすがる気持ちで実行に移す。いわゆる「目の前の川を跳んで」しまう。しかし、実はやりたくない職員は口では盛んに理想論を唱えながら、事柄のマイナス面を重大視して腰が引けてしまう。腰が引けたまま評価に取り組んでも、必ず出てくる言葉は、「これは余分な仕事だ。本来の仕事ができない。残業が増える。もうこんなことはやめてほしい。こんなことをして本当に変わるのか。ただやらされるだけに終わるのではないか。」であった。

そもそもこういう役所の組織的な病理現象は絶えずどこにでも存在するものであり、この「病理現象の絶えざる治癒」のためにも評価に取り組むのだとひたすら心の中で呼びかけるしかなかった。職員の自己変革力の存在を信じて、それが表に出てくるように仕向けること。そしてその自己変革力の存在を見つけたらその芽を徹底的に育てること。ここから改革を進めていくのだと自分に言い聞かせる日々であった。

職員が評価に対して持つ「なんとなくいやだ」

とか「こんなこと、したくないな」という抵抗感の中身を分析してみると、それは、面倒などと言う表層的なことではなく、職員が評価作業を通じて、意識構造や価値観の転換を強いられることに対する反作用と言えた。

1つ目は、「安定感から不安定感へ」である。自治体職員の意識構造の特徴としては、なんといっても安定志向である。ところが評価システムでは成果目標を明示することとなっているので、職員はその目標値が達成できるかどうか、これまでに体験したことのない不安定感を味わうようになる。職員にとってこの不安定感は面白くないのである。

2つ目は、「役所起点から生活者起点へ」である。職員は役所になじむに従い、自己本位、つまり役所起点の判断になってその思考経路が深部に記憶されている。ところが評価システムの導入で生活者起点の判断を強いられるので、職員にとってこれも戸惑いであり、甚だ不愉快なことなのである。

これらの背後には、長年の中央集権システムや右肩上がりの経済成長の中で育ってしまった県庁の組織文化を垣間見ることができた。

- ①大きな制度設計は専ら国の仕事で、県は国の出先機関的にそれをいかに勉強し、咀嚼して市町村の指導に当たるかということに終始し、そのテクニクのみが発達してしまう。
- ②国庫補助事業は全て無条件に善なりで、それをどれだけ多くこなすかが価値基準となり、補助金をもらうことが目的化してしまう。
- ③補助金や許認可をとることに熱心なあまり、「お上」の方ばかりに関心が向いて、結果を問うことなく予算獲得とその消化に良くも悪くも没頭する。
- ④予算は1円も残さず消化しなければいけないかのような状況をつくってしまう。予算残額がでると、来年の予算獲得に影響するといわれる。また、予算残額について監査等で説明しなければならないので、使い切る方を選ぶ。
- ⑤一般的に年功序列人事の風土の中で、大過のないことが価値基準となり、思考停止の前例踏襲のスタイルが蔓延する。

このような従来の県庁の組織文化を改革し、地方分権時代に向けて政策の自律性を高め、国への

追随的発想からの転換、成果志向、結果重視の行政運営を確立することを目指したのが評価システムであったが、この作業に巻き込まれた職員の大半の反応はかなり消極的で、評価作業が進めば進むほど抵抗感が大きくなっていく感じであった。そういう現実の厚い壁を思うとき、「結局無駄な作業」になるのではないかというところへ行きつ戻りつしながら「膨大な作業」になだれ込んでいったのだ。筆者は部下担当に「がまんするしかない。推進担当側が切れたら終わりだ。」とひたすら言い続けるしかなかった。

そのような時期に日経ビジネスの取材を受けた。1996年11月25日発行の『日経ビジネス』第867号の記事「市場原理を導入、“行革先進県”三重県の試み」が出て初めて大きな反応が起きた。すぐさま自治省（当時）から「取り組みを説明してほしい」旨の電話が入った。筆者はすぐには対応しなかった。喜び勇んで説明に行く気になれなかったのである。それでも放っておけないと思い直し1ヶ月後に部下を自治省へ説明に行かせた。また大手自動車会社の本社経営企画担当課長の視察来庁にびっくりしたこともあった。この時期を境にマスコミの取材と自治体職員等の視察が急増した。筆者らはその対応時は充実感に満ち勇気づけられることが多かったが、説明を終え一步廊下へ出ると、そこには冷たい空気が流れ出会う職員の視線は棘であった。マスコミに取上げられることによって職員が抵抗から受容へと態度を変えることはなかった。マスコミに取上げられるたびに、筆者たち推進側に対する反発意識がさらに高まったといつてよい。

この時期が、評価システムを推進する立場の事務局にとってつらく、不安な時期であった。もう戻ることは許されない。この膨大な作業が膨大な無駄に終わるはずはない。その理念を信じ徹底的に追い求めて行くしかない時期であった⁶。

(5) 企画課の抵抗

評価システムの構築を進めていくに従い、評価するためにはそもそもどういう目標（数値）があったのか、その目標が明確でない限り評価できないことが分かってきた。そこで当時、同時進行で

進められていた総合計画作りと密接に連携しなければ今後の評価作業もできないと考え、評価システムに基づいて総合計画に数値目標を設定していくことを知事に進言した。知事と企画部長と筆者の3人の前で、そのように検討するように知事から企画部長に申し渡された。「部内で検討します」ということで別れたが、いっこうに返事が帰ってこなかった。2週間後、企画部長から「検討の結果、従前どおり計画作りを進めたい」との返事が知事になされた。評価システムは未知のものであるというのがその理由であったが、そこには縦割り意識がありありとうかがえた。計画作りは自分たちの領域であり、評価システムとの連携などはわずらわしい、これまでのやり方で策定作業を進めるということで委託シンクタンクとも打ち合わせをしている。そういうことだった。筆者たちは理解されなかったことへの腹立たしさよりも自分たちの無力感に泣くしかなかった。部下たちからはため息が漏れた。評価システムのその後の発展をだれが確信できただろうか。

そのようにして計画作りはすすんだが、それから1年たったころ知事は「やはり数値目標をかかげるべきだ」と企画部に命じた。方針転換で時間的に相当な無理があったが作り上げられた。そのとき評価システムを皆がどれだけ意識していたかは、分らない。

1997年11月新しい総合計画「三重のくにつくり宣言」が公表された。これからの県政のあり方を、1人ひとりの住民を「生活者」としてとらえその自主性を尊重しながら支援していくことを行政の主たる目的とする「生活者起点の県政の推進」とし、具体的な政策体系は、5本の政策の柱、20本の政策、67本の施策で展開され、その下にツリー状に事業が体系化された。

行政使命、政策、施策、事務事業という政策体系ができ上がり、目的と手段の連鎖関係が体系的に整備され、評価システムとのリンクが見えやすい形となった。総合計画には、283の数値目標（1998年3月策定の実施計画の数値目標597を加えると880の数値目標）が掲げられた。その数値は、必ずしも評価システムの成果指標と全て整合しているわけではなかった。しかし、整合・不整合の精査は後々の課題として残しながらも、数値目標

を公表した意義は大きかった。

(6) 財政課の抵抗

本格導入が始まり、研修が進み、各課での評価表への記入が進んでいた頃、筆者は県庁の廊下で、財政課の中堅職員にばったり会った。これは良い機会だと思い、評価と予算編成との連携について、こちらから依頼するスタンスで話しかけた。やり取りするうちに、「理念的には分かるが財政課としては結局は使えない、使わない」ということであった。筆者はあれやこれやと必要性を訴えたが平行線で廊下での立ち話は1時間半に及んだ。予算の査定に使えないというその根底には自分たちの査定権限を奪われるのではないかという恐れがあったと思われた。

結果的には知事のリーダーシップにより予算要求書への評価表添付の義務付けは実現し、知事査定するとき手元に評価表を置きながら査定するという場面の写真は新聞に掲載された。しかし実際にどれだけ活用されたかは定かではなく、形だけに終わったようでもあった。筆者らはそれでも1歩前進だと思っていた。

筆者は次のような予算査定の場面を思い出す。筆者は要求側の立場で予算要求資料を前にして座っている。予算査定議論を聞いているうちに、自分の描いていた構想が音を立てて崩れ落ちていくのを感じ始める。「それならば、やらないほうがましだな」と頭の隅で考え始めたちょうどその頃、言い訳程度の形ばかりの予算がつき、「ありがとうございます」と言って頭を下げている自分がそこにいる。こういう「予算づけの中途半端化現象」はいつも頻繁に起きていた現象で、その背景には、要求どおりに認めることは査定側（財政当局）の敗北であるという意識があった。査定側はとにかく要望額を削ることに全知全能を集中させる。真に必要な額で要求した場合でも同じ扱いを受けるので、職員は翌年度からは「切り代」をつけて要求するのが当たり前となっていき、査定側はその「切り代」をいかに適正に探り当てるかが職人芸となっていく。この腹の探り合いが予算編成作業であり、この駆け引きを「政策議論」と思い込んできたのである。

このような予算査定では、政策体系における上位の目的への妥当性の議論は不要であり、またどれだけの成果をあげるかの有効性の議論も不要である。そこではいかに予算額を押さえ込むかという意味での効率性の議論が主にならざるをえないのである。

そもそも予算編成は、その時代の流れにふさわしい方式がとられてきた。

右肩上がりの時代が長く続いたが、この時代では、そもそも評価自体を本格的に実施する必要がなく、自治体は補助金を受ける、あるいは交付税で措置される起債をするなどして、できるだけ多くの事業を実施していくことがまちづくりであった。企画課は、財源が確保できたらという仮定の下に、綺麗な夢物語を1冊の総合計画書としてとりまとめ、右肩上がりの坂の上に仰ぎ見る美しい雲を追い求めた。財政課は、夢は夢として横に置きながら、現実の毎年の歳入見通しのなかで特に新規予算要求を中心に個別に査定作業を重ねた。

ときに財政危機の時期には、緊急避難的に予算規模縮小のため、各部課の既存予算に対する一律的、強制的な減分主義がとられた。そしてその召し上げた財源の配分の仕方は、もちろん従来の個別の増分主義査定であることに変わりなく、まさにスクラップ・アンド・ビルドの予算編成として定着していた。

さらに1991、2年のバブル経済崩壊後、財政状況が厳しくなればなるほどスクラップが強調された。予算編成方式は従来のまま、効率性の追求のみに傾斜していった。

筆者は、評価システムはそもそも政策課題解決のために何をすべきかを考える（＝ビルド）道具であり、そのなすべきことが明確になったときに、優先度の劣るもの、圧縮または廃止すべきものが見えてくる（＝スクラップ）という道筋を考えていた。すなわち「ビルド・アンド・スクラップ」である。

職員はみな、権限のある部局を注視し、その風になびく。当時の評価担当は何の権限もなく、ただ煩わしい作業を押し付けてくるところにすぎなかった。

財政課が実質的に評価表を使わない限り、評価システムは県庁内で真剣に取り組まれることは期

待しがたい。当初は予算との関係をそれほど強調せずに、意識改革と政策形成を狙いとしたが、日々の現実の職員の消極的反応を見るにつけ、予算編成とのリンクが喫緊の課題となった。そこで筆者は「財政課に使ってもらうために、財政課が使いやすいように様式を変更してもよい」と決断した。それがどんなものになろうとも、やむをえない。筆者は「これは政略結婚だ。戦略的妥協だ」といって、部下を説得した。財政課が了承し、両者間で様式変更についての検討が始まった。結果としては評価表の基本的な骨格を維持しながら、2年目（1997年度版）の評価表には、予算の種別欄、関連施策欄、基本事務事業欄、全体事業計画欄、根拠法令欄、事務事業種類欄、財源内訳欄、公共関与の妥当性評価欄、手段の妥当性評価欄、コスト節約のための問題点欄などが追加された。さらに3年目（1998年度版）からは特に新規事業については検討要素が不足しているとして新たに新規事務事業目的評価表を作成し、それまでの記入スペース2頁の制約から最大6頁まで柔軟に増減できる様式に変更した上で、総合行政課題欄、現状と課題欄、予想される副次的効果・派生的効果欄、総合計画の目標欄、実施上の課題の検討欄、参考事例等欄などが追加された。そして「予算見積書の添付資料としては原則評価表を添付し、重複する内容の資料は不要とする。また、企画課が所管する重要事業調書として活用する」旨の通知が出されるまでに至った。このようにして、評価表を財政課は予算査定に、企画課は総合計画の進行管理にそれぞれ活用しようとする姿勢が初めて明確になったのであった⁷。

（7）目的評価表の公表への抵抗

評価表の公表については、導入当初、将来公表するとは一言も言っていなかった。評価システムを実務として構築できるかどうかが最大の課題であった段階で、そこまで考えが及ばなかった。公表を思いついたのは、システムの意義を認め真面目に取り組む職員があまりにも少なかったからであった。反発する職員、無視する職員など、その姿勢は幹部職員を含めいろいろであった。さまざまな課題はあるが、次の段階として職員全体を本

気にさせる力、押し上げていく仕掛けは、もはや研修とか説明ではなく、公表しかない。公表まで突き進んでこそ、一応の区切りとなる。膨大な無駄であったかどうかはその時点で初めて判定されると考えた。もともと自己評価で出発したのであるから、その評価の客観性、妥当性を検証し確認する方法として、住民に公表して広く意見を求めることは必然的な一連の流れでもあった。

部長会議で筆者が提案すると「公表したら職員は本当のことを書かなくなる」、「公表するとまずいことがあり、今後自分の首を絞めることになる」、「目的評価表の記載内容が公表に耐えられない」、「そこまでしなくてもよいのではないか」等々の意見が続出した。強く反対したのは財政を握る総務部長であった。

筆者は「公表の決定をしないと職員は本気で書かないし、公表に耐えられる内容にならない」と主張したが、結局提案は採用されなかった。

公表に踏み切るまでには、同時に進んでいた「カラ出張問題」という県庁組織文化の大転換のドラマを克服することが必要であったのである⁸。

1996年1月知事が年頭訓示で評価システムについて熱く語っているとき、全国市民オンブズマン連絡会議は都道府県の監査委員事務局を対象として旅費、食糧費に関する情報開示請求を決定した。

評価システムの研修が盛んに行われている5月、阪神淡路大震災が発生した日（1995年1月17日）に三重県監査委員事務局職員が佐賀県へ新幹線に乗って2泊3日の出張をしていることが新聞で報道された。もう言い逃れはできなくなった。9月、知事はカラ出張による不正支出額は総計11億6,600万円余になり、それを課長級以上の職員で返還すると発表した。全国都道府県で3番目であった。これを機に官官接待を全廃するとともに、情報公開の範囲が拡大された。原則として全面公開とする開示基準規則を制定したのは当時としては全国で初めてであった。これまでの「非公開が原則、公開が例外」から「公開が原則、非公開が例外」へ流れが一気に加速したのである。ただし、それは「請求されたら開示する」という受身の範囲内でのことであった。

筆者の提案は、この受身の情報公開（第1段階）から、積極的な情報提供（第2段階）へ大きく1歩

踏み出すことを意味したが、第1段階の対応をきちんと踏み終えるまでは、第2段階には移れなかった。結局全庁的な合意となるまでには筆者の提案から1年間という時間を要したのであった⁹。

1998年2月27日、全ての事務事業目的評価表を公表した。公開対象は「1998年度当初予算用事務事業目的評価表」(3,381件)、公開場所は情報公開総合窓口で、分量はロッカー2段分であった¹⁰。

公表は恐れていた混乱をもたらすこともなく、「これまで知ることのなかった情報であり、公表は評価できる」「住民が県政の意思形成に参画できる第1歩である」と評価され、一種の驚きをもって迎えられた。反面、「分かりにくい」「もう少し説明がないと理解できない」との指摘とともに、公表の意義を過小評価する向きもあった。筆者は、公表による効果としては、まずは役所の組織文化の変革が大きいのと考えていた。住民は、情報非公開の時代、10万人、20万人の署名を集めて役所を攻めた。情報公開の時代は、ひとりの人間がじっと目を凝らしていれば、その「1刺し」で役所が倒れることがあり得る。そのことへの恐れが大きくなり、役所自らがその組織文化を変革せざるを得ないようになっていくのである。

首長も、議員も、民間事業者も、住民も「私は知らない」とは言えなくなり「読んでいないあなたが悪い」になってしまった。お互いに評価システムの公表によって生ずる自浄作用によって、筋の通らないことを押しつけにくくなった。また毎年の数値目標と実績の記録も経年的に残り、説明なく勝手に変更することも許されなくなった。こういう過程を重ねていくことによって議論がオープンな場で行われるようになり正論が多くなっていったのである。

情報公開は、官僚組織に自己を律することを強制する効果において、他のいかなる方法より優れている。この土台の変革なくして、評価の専門性を高めてみても実効性は望めないという思いが筆者には強かった。そして、行政の透明性を保ち説明責任を果たすだけでなく、政策情報、評価情報の住民との共有化による政策責任の共有化へ果てしなくステップアップしていくことを夢見たのである。

(8) 知事のリーダーシップ発揮

知事は、自らの改革理念の提示に始まり、それを受けてのさわやか運動と評価システム導入の英断以降、一貫して「本気と徹底」の姿勢を示し続けた。「相手が分かってくれないから、うまくいかない」と他者責任にして筆者らが無難なところにいようとすると、「それは、分かってくれない相手が悪いのではなく、分からせることが出来なかった者が悪い。相手が分かるまでやれ」と何度も返ってきた。そのように言われ続ける限り、うまくいかないのは自己責任だと自分に言い聞かせて努力するしかない。しかし、どうしても打開できない局面があった。そのときに知事のリーダーシップによる支えがなければ、評価システムは矮小化され、無視されていったらと思う。さわやか運動のキックオフ大会で「知事が要らなくなるくらいに評価システムをしっかりとやっていきましょう」と筆者が呼びかけたことは3の(2)で述べたが、「評価システムは知事を必要とした」のであった。評価システムがさわやか運動という行政改革運動全体の取り組みの根幹を担い、特に計画、予算という県庁の伝統的な基幹システムと一体となって機能するようにするためには知事のリーダーシップのすぐれた発揮が不可欠であったのである。

(9) おわりに

評価システムを立ち上げ、これをいかに実務の場で生きたものにするかに苦闘してきた筆者にとって、理論よりも「こだわり」が最初で、最後の拠りどころであった。「これでいいのか、役所は、地域は？」というこだわりである。

このこだわりから当たり前のことを行おうとする。が、必ず厚い壁にぶち当たる。思い悩む。もう止めようか。時にこの悩みを解き、思いを補強してくれる理論に出会う。元気づけられてまた実践に戻る。しかし、また壁に悩む。この繰り返しであった。

理論的成果を岬の灯台に例えるなら、時に波荒く、風をよけるためにはボートを灯台とは違う方向に漕がざるを得ないことがあった。灯台からは、

違う、違うと警告が発せられた。しかし、「ここで迷ってはいけない。自分の実感を大事にすべき」と考えた。灯台を目指しながらも、実務上はジグザグと漕いでいかなければボートは沈没してしまうのである。そのジグザグの戦略を誰も教えてはくれない。自分で風を読み、波を読まなければならない。自分にこだわりがないと、そこで迷路に誘い込まれてしまう。揺らいでも、揺らいでも、自分のこころの底に澱のように残るもの。地域のどこが、どう悪くて、それをどのように変えようとしたのか。そのためには、役所のどこが、どう悪くて、それをどのように変えようとしたのか。まずは、日頃のこれはちょっとおかしいという正直な実感に始まり、度重なるにつれて、どう考えてもやはり直さなくてはという思いへ高まっていく。そして、その思いは、根源的な原因となっている「ここ」をこの方向に変革しなければならないという改革のこだわりとして凝縮する。それが最初の出発点であり、かつ最後の拠りどころであった。

注記

- 1 北川候補456,676票、尾崎候補443,861票で、12,815票差であった。
- 2 北川知事は1995年4月から2003年4月までの2期8年就任した。
- 3 1996年本格導入のときの手引書で「三重県事務事業評価システム」にしたのが命名の経緯である（三重県 1996b）。
- 4 室長は副知事が兼務し、筆者は室次長で、以下課長補佐級2人、係長級2人計5人の小体制で出発した。さわやか運動の様々な取り組み（職員提案など）の展開に人手をとられ、評価システム主担当は1人であった。コンサルタントの提示する指導内容、シート案を基本にして、これに対する研修会参加者の行政現場からの意見を吸い上げつつ検討を行っていった。
- 5 三重県職員労働組合は、1998年、1999年、2000年に評価システム運用に関する職員アンケートを実施し、その結果を公表している（梅田 2002a：62）。
- 6 職員の評価システム受容の大きな最初のきっかけは、1998年2月の評価表の公表である。その受容は

「公表するのなら、やむを得ずやらざるを得ないと観念する」という受容であった。次の大きなきっかけは1999年に評価表作成責任者として所属長名を明示したことである。当時筆者は地域振興部長で急に部下所属長が真剣になったのでびっくりしたが、その受容は「自分の名前が明示されるのなら、真剣にやらざるを得ないと観念する」という受容であった。このような受容の基本的性格は、今なお続いていると筆者は考えている。すなわち受容の底流には、「好きならば簡単なものにしたい、理屈が立つならいっそなくしたい」という思いが絶えずあるのである。

- 7 評価表と予算編成との連携については、その後も試行が続き、評価表の様式もその時々の方考え方により見直しが行われた。
- 8 「カラ出張問題」は、官官接待をめぐる食糧費問題についての全国市民オンブズマン連絡会議の追求の延長線上であきらかになった（梅田 2002b：17-27）。
- 9 部長会議で公表すると決まってから数日後に、企画部長が筆者を訪ねてきて「本当に公開するの？」と問いかけてきた。「本当だ」とこたえろと、「では真剣に読んでチェックしなければならないな」と部屋を去った。その後姿を忘れることができないが、それは当時のほとんどの管理職員の思いと行動であった。公表が決まってから公表までの1ヶ月間がそれまでのなかで最も真剣に評価表が読まれ、議論された時期となり、かなりの修正が行われた。筆者はこのようになることを予想していたので、知事に「正式決定から公表まで少なくとも1ヶ月の期間をおくこと」を提案していた。
- 10 三重県のホームページによる評価表の公表は、1999年4月からである。

参考文献

- Osborne, David, and Gaebler, Ted (1992). *Reinventing Government*. (Reading, Mass:Addison-Wesley). (邦訳: (1995)『行政革命』、日本能率協会マネジメントセンター)
- 梅田次郎・竹内泰夫 (1999) 「三重県の事務事業評価システム」、斉藤達三編著『実践自治体政策評価』、ぎょうせい、91-151
- 梅田次郎 (2000a) 「政策評価導入の体験—事務事業評

- 価システムの導入に対する三重県庁内の組織的抵抗」、『公共政策：日本公共政策学会年報2000』
- 梅田次郎（2000b）「行政評価のアウトソーシング」、島田達巳編著『自治体のアウトソーシング戦略』、ぎょうせい、75-96
- 梅田次郎（2001）「業績測定型・三重県事務事業評価システムの発展過程と展望」、『日本評価研究』、1（2）：69-77
- 梅田次郎（2002a）「意識改革と政策形成：三重県庁における自治体組織運営の変革プロセス」、『公共政策研究』、（2）：55-69
- 梅田次郎（2002b）「当たり前役所をつくる「さわやか運動」物語」「行政評価の実践」、梅田次郎・福田志乃編『自治体実行主義』、ぎょうせい、1-54
- 梅田次郎・西城昭二（2003）「三重県における政策評価の取組み—事務事業評価システムからみえ政策評価システムへ」、古川俊一・星野芳昭・梅田次郎編『行政評価ゼミナール』、ぎょうせい、1101-1173
- 梅田次郎（2004）「行政評価の原点」、梅田次郎・小野達也・中泉拓也『行政評価と統計』、日本統計協会、21-69
- 三重県（1995）『さわやかニュース』創刊号
- 三重県（1996a）『三重県「さわやか運動」推進大綱』
- 三重県（1996b）『事務事業評価システムの手引き』
- （2008.1.23受理）

Hardships of the Pioneer : Introducing the Evaluation Systems in the Mie Prefectural Government

Jirou Umeda

JMAC Government and Corporate Governance Renovation Sector
jirou_umeda@jmac.co.jp

Abstract

Among the administrative reforms that Mie Prefecture has striven to implement since 1995, include unprecedented efforts towards introducing a project performance measurement system into core administration policies. Currently, this performance measurement system serves as the nation-wide model for evaluation systems being implemented by local governments. The author has been involved with these efforts in a position of responsibility in administrative affairs since their inception.

In this manuscript, by following the progress of the introduction of the project performance measurement system from within the Mie Prefectural Government, the author clarifies the conflicts that exist within bureaucratic organizations. Like the case of a government body that has employees who resist the introduction of such a system, while at the same time, having employees who make reform efforts on their own. The author also contends that the effectiveness of policy evaluation cannot be expected without the improvement of current conditions within bureaucratic organizations.

Keywords

decentralization reform, attitude reform, policy making,
performance measurement, bureaucratic organization

【研究論文：依頼原稿】

都道府県の評価の10年 — 定量評価のこれまでとこれから —

小野 達也

鳥取大学地域学部

onotatsu@rstu.jp

要 約

日本の公共部門における行政評価・政策評価の導入・普及において先導的な役割を果たしたのが都道府県である。現在では、すべての団体が種々の取り組みをしており、評価結果を活用する上で欠かせない客観性を確保するため、定量的な方法も様々行われている。しかし、定量評価の現状には、少なからぬ問題があることも明らかである。本稿では、都道府県における評価制度・システムの現状を踏まえ、評価の質や結果の活用に及ぼす影響が大きく、同時にまた定量的な方法一般の基礎に関わる3つの場面、①業績測定におけるベンチマーキング、②住民調査（満足度調査とCVM）、③公共事業の費用便益分析を取り上げ、典型的な問題を具体的に示し、解決のために何が必要かを考察する。

キーワード

都道府県、定量評価、業績測定、住民調査、費用便益分析

1. はじめに

日本の行政評価・政策評価は、まず都道府県において導入され、数年のうちにほぼすべての団体に行きわたった。都道府県の取り組みには先進的なものも多く、市町村や国の評価制度にも様々な影響を与えた事が知られている。

都道府県における評価においては、評価指標の算出や比較など定量的な方法を用いることで評価の客観性を確保しようとする試みも早くから種々になされてきた。しかし、その現状には、多岐にわたる問題があるのも事実である。

本稿は、まず、都道府県の評価制度・システムがどのようにして現在の姿に至ったかを整理し、

その上で、行政評価・政策評価が抱える諸課題のうちでも、将来の定着・進化の鍵を握ると考えられる定量評価について、特にその基本に関わる問題について考察するものである。

なお、都道府県では、どの団体も毎年のように評価制度・システムを改良したり、見直したりしており¹、今回あらためて全都道府県のホームページを通じ、行政評価・政策評価の最新状況（2007年10月時点）について定量的な取り扱いを中心に調査した（以下「今回の調査」と呼ぶ）²。

また、本稿では、広義・狭義政策や行政活動を評価する制度・システムを総称して「行政評価・政策評価」と呼ぶ³。しかし、本稿で取り上げる多くの問題は、評価制度・システムと呼ばれる仕

組みに限らず、個別分野における計画の目標管理などにも通底すると考えられる。

因みに、今回の調査によれば、自らの団体における評価活動の総称として「行政評価」という語を用いているのは16団体、「政策評価」は18団体であった⁴。

2. 都道府県が先行した日本の行政評価・政策評価

三重県が1996年4月に、一般に本邦初⁵の評価制度であるとされる事務事業評価システムを導入すると、その後を追うように多くの都道府県が続き、1999年9月末に自治省（当時）が初めて全国の自治体における取組状況⁶を調査した時点で、47都道府県のうち16団体が「既に導入済み」、10団体が「試行中」、21団体が「検討中」と回答している。この時点の市区町村（政令市を含む）では、全3,252団体のうち、既に導入済みであったのが2%弱（53団体）、試行中が1%強（45団体）に過ぎなかった（自治省 2000）。

その後の、自治省・総務省の調べ（自治省 2001、総務省自治行政局 2001, 2002b, 2003, 2004, 2006, 2007）もあわせ、自治体における取組状況の推移を図1に示す。都道府県が先行し⁷、政令市が後に続き、現在は政令市以外の市区町村にも広く普及した段階であるといえる。なお、市区町村の中で

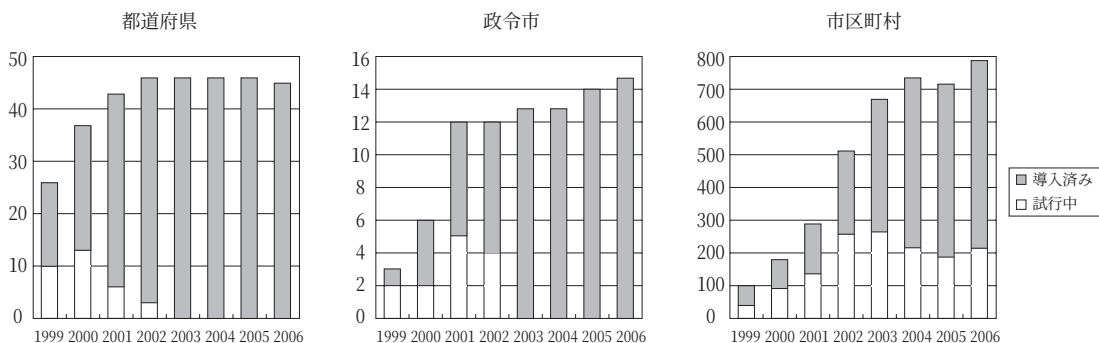
も、「導入済み」の割合は中核市が89%、特例市90%、その他の市区48%、町村16%となっており（2006年10月）、都道府県・指定都市を含め、団体の規模が大きいほど導入が早いという傾向も明らかである。

都道府県では、数の上で先行しただけでなく、注目を集め、他の自治体等に少なからぬ影響を与えたと思われる特徴的な取組も次々に登場した。三重県の事務事業評価システム（1996～）を筆頭に、静岡県業務棚卸表（1997～）、北海道の時のアセスメント（1997～）、岩手県の総合計画の進行管理（1999～）、青森県の政策マーケティング（1999～）、滋賀県のしがベンチマーク（2000～）などをその例として挙げることができよう。

自治体の行政評価・政策評価が、政策・施策・事務事業という階層のどのレベルを対象としているかを、総務省の調査に基づいて、表1に示す⁸。都道府県の特徴は、多くの団体が政策・施策レベルを対象にしていること、また都道府県以外ではほぼすべての取組が対象としている事務事業を対象としていない団体がいくつかあり、すべての事務事業を対象とする割合も都道府県以外の団体よりも低くなっていることである。

この背景には、行政評価・政策評価の取組がまずは事務事業を1つ1つチェックするという活動として広まり、その後必要性や評価能力に応じて、施策レベルや政策レベルの評価に拡張されてきたということや、全事務事業を一律に評価すること

図1 自治体における行政評価の取組状況（年度別・団体数）



（注）調査時点は1999年度が9月末、2000年度は8月末、2001～2004年度は7月末、2005年度は1月1日、2006年度は10月1日。

（出所）各年度の自治省・総務省「地方公共団体における行政評価の取組状況」をもとに筆者作成

表1 自治体における行政評価の対象（2006年度・割合（％））

	都道府県	政令市	中核市	特例市	市 区	町 村
政策	35.5	26.7	12.1	15.4	13.5	5.8
施策	84.4	66.7	42.4	38.5	36.5	18.5
事務事業	86.7	100.0	100.0	97.4	98.0	97.8
うち全事務事業	44.4	66.7	51.5	51.3	44.4	53.6

（注）表中の数字は、行政評価を導入済みまたは試行中の団体に占める割合（％）である。

（出所）総務省（2007）「地方公共団体における行政評価の取組状況」

表2 都道府県における評価対象の変遷（年度別・団体数）

	1999	2000	2001	2002	2003	2004	2005	2006
政策	12	11	14	16	18	18	17	16
施策	23	25	27	33	35	35	40	38
事務事業	38	40	41	44	45	46	46	39
うち全事務事業	18	15	17	19	20	21	22	20

（出所）自治省・総務省「地方公共団体における行政評価の取組状況」

の意義や負担が、都道府県と基礎自治体とは異なることがあると考えられる。

都道府県における評価対象がどのように変化してきたかを表2に示す。2005年度に、事務事業を対象とした評価を行っている団体が46、全事務事業については22あったが、06年度にはともに減少に転じたことがわかる。また、政策レベル・施策レベルもそれぞれ04年度、05年度以降わずかながら減少に転じており、評価対象を広げていくという動きが都道府県においては、一段落となっていることが窺われる。

さて、個々の都道府県の評価制度・システムの1つの典型は、総合計画の政策体系に沿って、階層の複数のレベルを評価対象にするものである。2006年度の総務省の調査では評価結果の活用方法として総合計画の進行管理を挙げている団体が29あり、また05年度の同調査で、行政評価が総合計画と連携しているとした団体が31ある。

今回の調査によれば、評価対象が総合計画の体系に基本的に対応している団体が39あり、その中から、事務事業評価型の評価シートに総合計画上の位置づけを記入する欄があるだけといった形式的な対応のケースを除いて、総合計画の体系に沿

った評価が中心的な制度となっている団体が33あった。さらに、総合計画の政策体系のうち複数の階層を評価対象としている団体が25あり、複数の階層の評価を評価シートを用いて（単に評価指標を設定しているに留まらず）評価している団体が18ある（予定段階の団体を除く）。（表3）

また総合計画の進行管理を行う評価のほぼすべてが、評価指標を設定し、評価シートの中でも指標に重きを置く業績測定型となっている。吉川（2004）は、米国の地方政府の業績測定を、地方政府全体の大括りな課題・分野別に設定した地域全体に関わるアウトカム指標の目標達成度を重視

表3 都道府県の評価制度・システムと総合計画の関係

タイプ	団体数
評価対象が総合計画の体系に対応	39
総合計画の評価が骨格	33
総合計画の複数階層を評価	25
総合計画の複数階層を「数値指標を含む評価シート」で評価	18
オレゴン型業績測定	11
テキサス型業績測定	10
オレゴン+テキサス複合型業績測定	8

（出所）筆者の今回の調査による。

し、戦略計画の目標管理を行うコミュニティ型業績測定（オレゴン型）と、各部局の施策・事業レベルで、アウトカムの他にアウトプットやインプット、効率も重視し、予算サイクルにフィードバックする部局経営型業績測定（テキサス型）とに分類している。これにならって都道府県の総合計画ベースの評価を分類してみると、オレゴン型（大括りのアウトカム指標を重視して、総合計画の進行管理を主目的とするタイプ）の取り組みが、総合計画の評価を骨格とする33団体のうち11団体、テキサス型（施策・事業レベルでアウトカムに加えてアウトプット、インプット、効率等に着目し、総合計画の枠内で予算要求や査定への活用を重視するタイプ）が10団体、上位階層のオレゴン型と下位階層のテキサス型を合わせ持つ複合型が8団体である¹⁰（表3）。なお、総合計画からほぼ独立した事務事業評価型の評価の仕組みを持っているのはわずか3団体であった。

3. 都道府県における公共事業評価

ところで、自治体の行政評価・政策評価に関し、あまり表立って論じられないのが公共事業の評価である。いうまでもなく、公共事業はほとんどの都道府県にとって、予算規模からいっても、本来果たすべき役割からいっても重要である。しかし、公共事業の評価では費用便益分析など、公共事業以外の一般事業の評価にほとんど用いられない手法が必要なこと、公共事業の評価は大量の技術情報を踏まえてなされねばならないこと、さらには大規模な公共事業はすべて国の補助事業であり、国による公共事業の事前評価及び再評価（事業採択後一定期間を経過した事業を評価し、必要な見直しを行ったり、休止・中止したりするもの）の一環として都道府県が行う評価資料作成は国のPDCAサイクルに寄与することはあっても、（公共事業以外の補助事業のように）自治体のPDCAサイクルには直結しない仕組みであることなどのため、公共事業の評価と一般の政策を対象とした評価制度・システムの連携がほとんど見られないのが現状である。表4に、今回の調査結果を示したが、全体の評価制度・システムの中に公共事業

評価を位置づけていたり、全体の評価を所管する部局が公共事業評価も担当したりするなど、両者に何らかの関係が見られる団体は13であった。

日本における公共事業評価は、1997年12月に内閣総理大臣から公共事業関係6省庁の大臣に対して、再評価システムの導入と事業採択段階における費用対効果分析¹¹の活用について指示が行われたのを契機に大きく普及・進展した¹²。補助事業についていえば、この後、関係省庁から事前評価・再評価における都道府県の役割に関する通達が次々に出されている。

再評価では、都道府県が第三者機関の意見を踏まえて「対応方針」（再評価結果）を作成・提出し、本省庁はこれを「尊重」して、国庫補助の継続等に関する判断を行う仕組みである。事前評価については、従来部分的に行われていたが、より本格的に行うこととなり、都道府県が費用便益分析の結果などの「新規採択時評価に係る資料」を作成・提出し、所管省庁がこれに「検討を加え」、予算化に係る方針を決定する仕組みである。

このように、都道府県は各省庁の事前評価・再評価の作業に組み込まれているため、事業主体として自主的かつ独立に評価を行おうとする場合、義務として行う「補助金を獲得するための」評価作業との関係が微妙となる。このような状況で、事前評価・再評価として独自の評価を行っているのは、23団体・20団体であった（今回の調査で独自の評価基準や評価表を確認できたもののみ）。

また、本来必須であるはずの事後評価については、少し遅れて導入されている。1999年に関係省庁において試行の申し合わせがなされ、例えば国

表4 都道府県における公共事業評価への取り組み

取り組み	団体数
一般事業の評価と何らかの関係あり	13
独自の評価制度（事前評価）	23
" （再評価）	20
" （事後評価）…試行中を含む	9
第三者による評価（事前評価）	10
" （再評価）	46
" （事後評価）…試行中を含む	12

（注）独自の評価制度の団体数は、基準や評価表などを確認できたもの。

（出所）筆者の今回の調査による。

土交通省では直轄の事業について2003年度から実施されている。都道府県でも、今回の調査では9団体において具体的な取り組み（うち2団体は試行中）を確認できた¹³。

再評価にあたっては、学識経験者などからなる第三者機関の意見を聴取することが義務付けられているが、事前評価・事後評価についても、それぞれ10、12団体が行っていた。

4. 定量評価の重要性と諸問題

現在行政評価・政策評価を導入している都道府県は少なくとも5年以上の経験を有している。多くの団体は、評価制度・システムのバージョン・アップをすでに一度ならず行っている。また、都道府県以外にも早期に導入した団体は同様の実績を積んできている。これまで公共部門に存在しなかった評価の導入・定着に一定の時間を要するのは当然であるが、そろそろ行政評価・政策評価の成果や費用対効果を検証・評価してもよい時期になったともいえる。実際、最近では自治体におけるこれまでの成果を検証するために、アンケートや計量分析などにより評価制度・システム導入の効果測定を試みる論者が続々出てきている。

いうまでもなく、評価とは結果を生産することが目的の営みではなく、結果を活用してこそ意味のある企てとなる。もちろん評価シートの記入や顧客満足度調査などの評価作業自体を通じて意識改革がなされるといった効果もあろうが、それはあくまでも副次的なものであろう。

どのようなアプローチであれ、行政評価・政策評価の最大の目的は、行政資源の配分や政策・施策・事業の改善に資する情報を意思決定者に提供することであろう。また、アカウンタビリティの遂行をそれに次ぐ大きな目的として挙げることもできる。実際、国の政策評価制度を始め、ほぼすべての行政評価・政策評価の取り組みが、これらを目的として掲げているはずである。

すなわち、評価制度・システムの成否は、このような目的で活用できるような情報を提供できるか否か次第である。そのための必要条件としては、事実及び事実に基づく推論という客観性がまず求

められる。さらに評価作業の基本は比べることである。時系列で改善したのか否か、他の事業あるいは自治体と比べて良いのか悪いのか、目標などの基準を達成しているのか否か、事業を実施しない場合と比べてどれだけの効果があったのか、費用に比べて十分な便益が見込めるのか、など。

この客観性と比較可能性という2つの条件を満たすためには、数・量による把握すなわち計数・計量あるいは定量的分析が必須であることは明らかである。中央府省の政策評価制度を法的に規定する「行政機関が行う政策の評価に関する法律」の第3条第2項第1号には「政策効果は、政策の特性に応じた合理的な手法を用い、できる限り定量的に把握すること。」とあり、評価法を受けて閣議決定された「政策評価に関する基本方針」でも「できる限り政策効果を定量的に把握する」とされているが、このような考え方自体は多くの自治体にも浸透していると考えられる。

実際、都道府県におけるこれまでの行政評価・政策評価をみても、先陣を切った三重県の事務事業評価システムはその検討の過程で目標の数値化や達成度の測定が重視されるようになり（梅田2004）、事務事業評価型評価のモデルとなった。また、先に典型的な取り組みとして挙げた静岡県業務棚卸表、岩手県の政策評価、青森県の政策マーケティング、しがベンチマークなどは何れも数値データ、数値指標をベースに評価を行うものである。そして現在主流をなす、総合計画の体系に沿った評価は、そのどれもが評価指標を中心に据えるか、少なくとも重視するものである。都道府県における行政評価・政策評価の広がりや、NPM（ニューパブリックマネジメント）等の世界の公共部門改革という文脈から見れば、欧米流の業績測定が日本に渡来し、普及した過程ともいえ、数量情報の重視は当然の帰結であるともいえる。

一方、業績測定とは系譜を異にする公共事業評価も、近年の費用便益分析を重視するという流れは、事前評価・再評価・事後評価の何れのタイミングにおいても、数量情報に基づく客観的な議論と判断を求めるものである。

このように、都道府県における行政評価・政策評価には、評価指標などの定量的方法がすでにか

なり浸透しているのであるが、そのことによって評価結果の客観性が増し、評価結果がより活用されるようになってきたのであろうか。これについては、今のところ楽観は禁物であると考える。

様々な評価作業や評価結果を見ると、数量データの取り扱いに疑問を持つことが少なくない。本稿の後半では、①業績測定におけるベンチマーキング、②住民対象のアンケート、③公共事業の費用便益分析という3つの場面を取り上げ、特に基本的な問題について具体的に考察することとする。

5. 業績測定におけるベンチマーキングの課題

総合計画の指標に基づく進行管理や、数値指標を重視する事務事業評価、住民参画型のベンチマーキングなど、業績測定型の評価において主役を務める評価指標については、これまで数々のアンケート等において、数値指標の設定などの取り扱いが難しい、という現場の声が報告されてきた(例えば小野 2001、白川・小野 2002、総務省自治行政局 2002a、佐藤 2004など参照)。数値の評価指標の取り扱い上の課題といえば多岐にわたるが(例えば小野 2004、2006、2007、梅田・小野・中泉 2004など参照)、ここでは指標の目標値を設定して行うベンチマーキングに着目する。指標値の変化について掘り下げた分析を行わないことが前提の業績測定型の評価においては、指標値の改善状況を判断する手だてとして、目標値設定と達成度評価は、最も有効な方法であろう。

今回の調査によれば、行政評価・政策評価に取り組む45団体のうち44団体が数値の指標を用いた業績測定を行っており、そのすべてが数値の目標値¹⁴を設定している¹⁵。また、目標達成度¹⁶を数値で算出するのは24団体であった(表5)。

また、評価対象の政策・施策・事務事業を目標の達成状況に従って3～12段階に評定している(その多くは、総合計画の進捗状況を、達成度別の政策・施策・事業数の集計によって説明する)団体が28ある。その内訳は、数値化した達成度の水準によって直接評定する団体が14、数値化した

表5 都道府県の業績測定・目標管理への取組状況(その1)

取り組み等	団体数
行政評価・政策評価	45
指標を用いた業績測定	44
目標値の設定	44
目標達成度の数値化	24

(出所) 筆者の今回の調査による。

表6 都道府県の業績測定・目標管理への取組状況(その2)

取り組み等	団体数
目標達成状況に基づく評定と集計	28
計算した達成度に基づく定量的な評定	14
計算した達成度に基づく半定量的な評定	5
計算した達成度に基づく定性的な評定	2
達成度計算なしの半定量的な評定	6
達成度計算なしの定性的な評定	1

(出所) 筆者の今回の調査による。

達成度をもとに「予定を上回っている」「予定より遅れている」「予定より著しく遅れている」などと「半定量的に」評定する団体が5、数値化した達成度をもとに全く定性的に評定する団体が2あった。また達成度を数値化しないながら、実績が目標値を上回っているか否か(この部分は定量的に判断)を含め順調か否かを半定量的に評定する団体が6、達成度を数値化せずに定性的な評定を行う団体が1ある(表6)。

評価指標と目標値を数値で明らかにすることが望ましいことはいうまでもなく、その上で、実績値に基づいて達成状況を定量的に把握することが望ましいことも明らかである。ただし、数値化すればよしとする訳にはいかないのも当然であり、評価指標の設定から達成状況の把握・集計までの過程で、妥当性がどれだけ確保されているかが問われる。このような、目標値に対する達成度評価の妥当性については、次の通り、大小あわせて10の観点から、吟味することができると考えられる。都道府県の業績測定の現状を見てみよう。

①評価指標自体の妥当性

業績測定においても、社会科学一般における測

定同様、妥当性（測定すべきものを測定していること）がまず第一に求められる。多くの米国の業績測定に関する文献も当然そのことに触れている（例えばWholey 1999）。

都道府県において設定した指標の妥当性はどうだろうか。その手がかりは評価指標の説明情報にあるはずであるが、指標と政策等のロジックとの関係など、評価指標の妥当性に関する記述や吟味があるのが、指標を設定しているうちの6割に満たない25団体であった（表7）¹⁷。残りの19団体では、その種の説明が評価調書やレポートに見あらず、総合計画の中で指標を設定している場合には計画書の中にも発見できなかった。公表資料の中になくても、内部的に整理されている可能性もあるが、多くは期待できないであろう。今回の調査で、説明があるとカウントした団体でも、ごく断片的な記述で済ませているものが多く、政策等のロジックを踏まえての設定、関係者間の合意などによる妥当性の確保については、公表情報を見る限り、心もとない状況の団体が多い。

②目標値の性格と設定根拠

目標値がどのような性格をもつかによって、目標の達成、未達の評価は当然変わってくる。目標値の性格として、例えば、達成可能か否かは別として尽力しようという理想値、十分達成可能な期待値、最低限達成すべき限界値という3種類（田辺 2001）があったとして、目標値を達成できなかった場合の解釈は大きく異なるであろう。

また、目標値の数字自体には根拠が必要である。過去の実績や関連データ・情報に基づくのが基本であるが、それをベースに関係者間の交渉によって

決まる場合もあろう。何れにせよ、H.P.ハトリも強調しているように、その根拠は明確にする必要がある。つまり未達によって評価が低くなることを避けるために低い目標値を設定したり、逆に予算増を求めて高い目標値を設定したりすることが常に起こりうるからである（Hatry 1999）¹⁸。

ところが、都道府県の状況をみると、目標値の根拠と性格について、何らかの記述が見られる団体は25にとどまる（表7）¹⁹。しかも、困難度の3段階表示などシンプルなものが多い。

③フロー的指標とストック的指標の区別

フロー、ストックとは経済統計の用語として認知されているが、より一般に統計データの時間的属性（上田 2003）と考えられる。一般に、フローとは一定期間に発生する量、ストックはある時点に存在する量と定義されるが、その両者の間には、

期末のストック

＝期首のストック＋期中のフロー

という関係があると前提されている。

ここで、ある目的で講習会を開催し一定の技能を習得してもらうという事業を考えよう。1年間の参加者数はフローを測る指標であり、累計の参加者数（または参加したことのある人すなわち技能を習得した人の割合）はストックを測る指標である。フロー指標は、行政の注力に応じて値が上下するのに対し、ストック指標は、一定の注力を続ければ確実に上昇し、注力を止めても横這いとなる。問題となるのは、A事業の評価をフロー指標で、B事業の評価をストック指標で行い、その改善度や目標達成度を比べるケースである。そのまま比較することができないことは明らかであるが²⁰、今回の調査で確認したところ、少なくとも10団体において、フロー指標とストック指標を区別せずに目標達成度を計算して評定し、その結果を集計している。他にも、目標達成度を計算して評定・集計している団体では、フロー指標とストック指標が混在していれば、同様の事態が起こっていることになる²¹。なお、今回の調査で、フローとストックの分類をしていることが確認できたのは3団体（神奈川、岡山、愛媛）であった²²。

なお、評価指標には、上記の定義のフロー、ス

表7 都道府県の業績測定・目標管理への取組状況（その3）

取組み等		団体数
評価指標の妥当性に関する記述、吟味		25
目標値の根拠や性格の明示		25
目標達成に影響を及ぼす外部要因の把握		6
目標達成度の計算	フローとストックの区別	3
	等差型以外の増減の想定	(0)
	達成度計算式の使い分け	5

（出所）筆者の今回の調査による。

トックの何れにも当てはまらないものも多い。ある時点での状態を測るような指標、量の多少ではなく水準の高低で測るような指標などである。また、ストックの指標であっても、社会経済状況に起因する減少を防ぐ目的の事業である（一定の注力は現状維持に相当）場合もある。この点について、評価指標の時間的属性は、評価対象である政策等との関係という観点から分類すればよいだろう。すなわち、通常のフロー・ストック概念を拡張して、同じの量（質的側面も含む活動量）の政策等を実行することが指標値の維持をもたらすのが自然であればフロー的、指標値の上昇をもたらすのが自然であればストック的とする。このようにすれば、もとの定義であればフロー、ストックのどちらでもない指標も基本的に分類できる。

各指標がフロー的であるかストック的であるか分類したとして、望ましい取り扱いとしては、ストックの指標をフロー的指標に変換する（差分を計算する）、フロー的指標をストック的指標に変換する（累計を計算する）、両者の達成度は別に計算・集計する、などが考えられる。

④指標値の増減パターンの想定

t_0 年に計画の最終年度である n 年後 t_n 年の目標値 $g(t_n)$ を設定したとして、中間年 t_i における指標の実績値 $a(t_i)$ の達成度はどのように評価できるだろうか。最終年度の目標値と当該年度の実績値をそのまま比べる（あるいは目標達成度を計算する）という団体もあるが、多くの団体においては、中間年の目標値 $g(t_i)$ を明示的に設定したり、達成度の計算の過程で $g(t_i)$ に相当する量を求めたりしてから $a(t_i)$ の達成度を計算・解釈している。

この場合、評価対象の施策等の年度ごとの事情から $g(t_i)$ を設定する場合もあるだろうが、多くの場合（達成度計算の過程では、おそらくすべてのケースにおいて）、目標値まで「一定の」ペースを想定して次のように按分しているようである。

$$g(t_i) = (1 - \frac{i}{n})a(t_0) + \frac{i}{n}g(t_n)$$

ここで $a(t_0)$ は基準年における実績、基準値である。ここでは、一定のペースとして、グラフ上で

直線上に推移する、つまり等差的な軌跡を描くこととなっているが、機械的にこのことを前提することは適切であろうか。等差的に推移するということは、増加率は一定でなく漸減していくことであるが、GDPのように一定のペースといえは定率で、等比的に上昇することが自然な指標も多いはずである。この場合、中間年の目標値は次のように設定されるべきであろう。

$$g(t_i) = a(t_0)^{1 - \frac{i}{n}} g(t_n)^{\frac{i}{n}}$$

例えば基準値100から10年後の目標値400に対し、等差的な推移を想定すれば5年後の目標値は250となろうし、等比的であれば200となろう。

最終年度の目標値に対して中間年の実績をそのまま対照させる場合にも、等差的な推移を前提として解釈する場合が多そうである。なお、数年間の目標値・実績値の推移をもとに、中期的な最終目標値を設定する場合にも、どちらの軌跡を想定するかで、大きく変わることはいうまでもない。

このような指標値の増減パターンについて、何らかの判断がなされていると確認できた団体は1つもなかった（表7）。

⑤達成度計算式の使い分け

目標達成度の計算方法は、何に対する達成度を計算するかによって、2通りに大別できる。まず文字通り目標値そのものと比べるもので、目標年度（期末）の目標値 $g(t_n)$ と実績値 $a(t_n)$ とから、次の量を計算する。本稿ではこれを「期末型」の目標達成度と呼ぶ。

$$\frac{a(t_n)}{g(t_n)}$$

もう一つの方法は、基準値から目標値までの、目標達成に必要な変化量に対する達成度を計算するもので、次の量によって「期首の基準値からどれだけ期末の目標値に近づくことができたか」を明らかにする。本稿ではこれを「期間型」の目標達成度²³と呼ぶ。

$$\frac{a(t_n) - a(t_0)}{g(t_n) - a(t_0)}$$

なお、上では目標年度における計算式を示したが、中間年の実績 $a(t_i)$ の目標達成度を、最終年度の目標値 $g(t_n)$ あるいは中間年の目標値 $g(t_i)$ に対して求める場合も、同様に分類できる。

また、期間型について、次のように前年からの1年間について計算するのも、そのヴァリエーションの1つであるが、この計算をストック的指標について行くと、フロー的指標に変換して、その期末型の達成度を求めていることになる。

$$\frac{a(t_i) - a(t_{i-1})}{g(t_i) - a(t_{i-1})}$$

これは一律どちらが望ましいということではなく、目標値の性格などによって、どちらの達成度がより明確な意味を持つかが変わってくるであろう。例えば目標値が最低限達成すべき限界値であれば、期首の水準がどうであったかに関わらず、期末の時点での達成状況をこそ問うべきであろう。一方、達成可能か否かは別として掲げている理想値であれば、期首から期末にかけて、どれだけ理想に近づくことができたかを測るべきだろう。

したがって、リストにある評価指標のすべてが同じ性格の目標値でない限り、達成度の計算式は使い分けたほうがよいことになるが、今度は、異なる方法で計算した達成度の大きさを直接比べることができないという問題に直面することになる。この点について、唯一の解決策は、両方の計算をすべての指標について行い、個々の施策等の達成状況についての解釈・分析を、2種類の達成度の役割を踏まえつつ行うことであろう。

さて、目標達成度を計算している24団体がどのような計算式を用いているかを見ると、統一的に期末型の計算をしているのが11団体、統一的に期間型の計算をしているのが8団体で、何らかの形で2種類の計算をしているのが5団体であった。またこの5団体のうち、4団体は2種類の計算式の使い分けで、1団体だけが必ず2種類の計算を行う形

になっている²⁴。

⑥達成度計算における例外的ケースの処理

達成度を計算する場合、次のようなケースでは例外的な取り扱いが必要となる。期末型の達成度計算においては、目標値がゼロの場合、分母がゼロとなり計算できない。また指標の値を小さくするべく目標を設定する場合（例えば事故や苦情の件数など）、そのままでは目標に近づき、超えるに従って「達成度」は小さくなってしまふ。また期間型では、現状維持を目標とする場合に分母がゼロとなり計算ができない。また、本来は上昇させたいが下降が避けられず、現状よりも低い数字を目標とする場合（または、減少させたいが増加は避けられず、現状を上回る数字を目標とする場合）には、そもそも期間型の達成度を計算する意味がないだろう。

期末型、期間型のどちらかで分母がゼロとなる場合に、その指標だけもう一つの計算式に替えるケースがあるが、これは適切な取り扱いとはいえない。また期末型で小さい方がよい目標の達成度を計算する場合に、分子と分母をひっくり返して

$$\frac{g(t_n)}{a(t_n)}$$

を計算したり、次の量

$$1 - \frac{a(t_n) - g(t_n)}{g(t_n)} = 2 - \frac{a(t_n)}{g(t_n)}$$

を計算したりする例が複数の団体で見られるが、これらは100%の近傍でそれらしい値（例えば110%は91%、90%にそれぞれ「変換」される）になるものの、どのような根拠があるのか不明であり、それゆえ100%から大きく離れる場合に解釈に苦しむ値を取る。この場合には、シンプルに目標値から何%ポイント離れているかに注目すればよい。例えば、目標値が50で実績値が40であれば計算結果は80%、すなわち目標より20%ポイント上回っているものであり、実績値が60であれば計算結果は120%となり、目標に20%ポイント未達であることになる。

⑦外部要因の把握

アウトカム指標、とくにロジック・モデルの最終段階に近いアウトカム指標の推移には、当該プログラム以外に、他の公共部門の政策や景気などの経済社会情勢、地域の企業や住民活動などの外部要因が影響を及ぼすことがあるのはいうまでもない。これは①で述べた指標の妥当性の問題でもあるが、指標を設定して算出した後に、実際に測定したアウトカムの値やその変化に、外部要因がどれだけ寄与しているかは、実績値の解釈や目標達成度の評価を大きく左右する可能性がある。

業績測定は外部要因の効果を取り除いた正味の効果（インパクト）を突き止めようとするものではないが、目標達成度を解釈する上で、外部要因を全く考慮に入れないのでは、大きく現実を見誤る可能性がある。都道府県の状況を見ると、何らかの形で外部要因を検討する仕組みが見られるのは6団体にとどまっている（表7）。

⑧達成度の解釈、達成状況の分析

目標達成度を何のために計算するのか。あるいは目標値と実績値を何のために比べるのか。それはPDCAサイクルでいえば“C”の手段であり、次の“A”、“P”に繋げるためにある。続く⑨に述べるような評価や、⑩で述べるような全体の集計をすること自体が主目的ではないはずだ。

目標に届かなかったとすれば、それは何を意味するのか、また何故なのか。指標の妥当性や目標値の性格、外部要因の働きなどを踏まえて分析や解釈がなされるべきであろう。同時に何が言えないのか、何が不明なのかも検討されねばならない。

ところが、目標値を設定している44団体の状況を見ると、個々の指標の目標達成状況について、何らかの解釈や分析が示されているということですら、28団体しか該当しない。残りの16団体はただ数字を示しているのみである。解釈や分析をしないまま数字を独り歩きさせることの可能性は、アウトカムのモニタリングにおける最大の陥穽である（Affholter 1994）といってもよい。

⑨達成度の評価

目標の達成状況を3段階や5段階に評価するということは、結果をわかりやすく要約するために有

効な手だてである。しかし、評価をするからには、できるだけ明確で、どの指標についても等しい基準に基づいてなされることが望ましい。

できるだけ明確な基準ということであれば、定性的な評価よりも半定量的な評価が望ましく、達成度の数値に応じて自動的に定まる定量的な評価が最も望ましい。しかし、どの指標についても同じ基準ということを考えると、定量的な評価を行うためには、③～⑥の観点に照らして、異質な指標や目標達成度が並ぶことがないようにする必要がある。指標や目標達成度の性格が異なるのであれば、また①②の観点が不十分であれば、達成度の計算や目標値・実績値の比較は危ういものとなり、評価結果は一層曖昧なものとなる。

⑩達成度の評価結果の集計

達成度の評価を行う団体の多くが、A評価のプログラムが何%、B評価が何%といった要領で、総合計画全体の進捗状況を達成度の評価段階別の政策・施策・事業数の集計によって説明している。

評価の段階別に集計するということは、評価結果を相互に比べることと同値であり、互いに比較できること、即ち同一の基準で評価されていること（⑨で述べた観点）が前提である。評価にいたる過程で、統一的な取り扱いをしないのであれば、2またはそれ以上のグループに分け、それぞれのグループごとに総括してはどうだろうか。

また、少なくとも3つ以上の団体において、全指標の評定の平均点を計算している。しかし、これら3団体を含めほとんどの団体の評価結果は間隔尺度の数字とは見なせず、順序尺度であるからには、平均を計算できないことを指摘せねばならない²⁵。目標達成度を計算すれば、その結果は基本的に比例尺度であり、評価とは、情報をわかりやすく要約するために情報量を削って順序尺度に水準を落としているのである。比例尺度の情報の比較可能性に自信が無く、その問題への対処として数字の水準を落とすことにはある種の妥当性があるといえようが、そこから平均を計算して勝手に間隔尺度として扱うことはできない。順序尺度の評価の結果を集約するとすれば、分布をグラフなどによって可視化すればよい。

さて、①～⑩の観点について順に眺めてきたが、以上を踏まえると、例えば「わが県の基本計画（計画期間5年）で掲げている政策目標50本の、2年目終了時点の達成状況は、満足すべきA評定（最終目標に対する達成率50%以上）が8割以上であり、きわめて良好です」という表明があった場合、それが客観的な評価結果の妥当な総括である場合もあれば、明確な根拠を欠く、ほとんど意味のない言説である可能性もあるといえる。

J.S.ホーリーも指摘している通り、業績測定システムを評価する基準は、前提条件としての技術的な質と、業績情報が活用されているか否かという2点であるといつてよい（Wholey 1999）。後者はそもそも業績測定の目的であるから当然であるが、前者は前提条件であり、それが満たされなければ、活用できないし、活用しようとすればおかしいことになる。すなわち、個々の政策・施策・事務事業のベンチマーキングにおいて、目標達成度情報が信用できるものでなければ、活用のしようがない。現場で活用できていないケースにおいては、当事者が自ら設定した評価指標・目標値を信用していない状態もあると思われる。

業績測定や、数値の評価指標の役割に過剰な期待を抱いてはならないし、その限界はわきまなければならない。しかし、そのことと、評価指標等の数字の取り扱いが不十分であることに起因する諸々の問題とは、峻別しなければならない。

6. 住民調査の課題－満足度とCVM

今回の調査によれば、少なくとも16団体以上が、業績測定型の評価に用いるデータの収集手段として、住民調査を積極的に活用していることがわかった。典型的なアプローチとしては、個々の政策・施策・事務事業やより具体的な施設・サービスについて満足度などの形で評価してもらい評価指標のデータとする、政策や施策のリストについて満足度と重要度を尋ね、重要度（平均点）と満足度（平均点）の乖離の大きさをもって行政ニーズを測る、などである。

まず、一般のアンケート同様、標本調査の基本を踏まえる必要があることはいうまでもない。例

えば、標本誤差の見積もり、調査票回収時に発生する偏りのチェックなどが必須であるが、公表資料を見る限り、このような事が明確になっている例は少ない。また、無作為抽出の標本と有意抽出の標本を、あるいは異なる抽出方法の標本を無造作に混合してしまう例も見られる。

また、標本の層別抽出を行うにあたり、各層から同じ大きさの標本を抽出する（例えば、各年齢階級から同じ人数を選ぶなど）ことがしばしば行われる。この方法は、層ごとの精度（標本誤差）をほぼ均一にするという意味（例えば、どの年齢階級についても同じ程度の精度の結果が得られる）の利点があることはもちろんであるが、全体の結果を推定する際は当然抽出率に応じた加重平均を計算する必要があるにも関わらず、単純に集計してしまう例がある（不偏推定にならない）。

次に、満足の程度を「満足」「やや満足」など3～5段階で尋ね（重要度について尋ねる場合も同様）、回答を数値化して（あるいは始めから数字を示しておき）、平均を計算することがよく行われる。この場合、本来間隔尺度のデータでなければ平均は計算できないから、調査票に予め5～1、+2～-2などの数字を付して等間隔の5段階評価であることを意識してもらうなど、できるだけ間隔尺度としてのデータを得るようにする必要がある。2005年度までの宮城県の県民満足度調査のように0～100のスケールで回答を得れば、これは比例尺度のデータとして平均はもちろん、変化率を計算するなど各種の取り扱いが容易になるが、例えば数十本の政策等について記入してもらうとすれば記入者の負担もまた大きくなる²⁶。

上記のように計算した平均点や「満足」「やや満足」の回答割合など、満足度を数値化した場合、その時系列の推移に関し解釈に苦しむことがある。即ち、施策を充実させたはずなのに、満足度が低下してしまうようなケースである。そもそもこの種のアンケートで測ることができるのは満足度の絶対的な水準ではなく、相対的な評価であることを確認しておく必要がある。質問は絶対水準を尋ねる形式であっても、回答者の多くは要求水準との比較、他者との比較、他の施策との比較など相対的に評価することは明らかであろう²⁷。

政策・施策など項目別の満足度（あるいは充足

度など）及び重要度の得点（回答の平均など）を2次元にプロットし、重要度が高く満足度が低い項目はニーズが大きいとするような分析も広く行われている。民間企業の顧客満足度調査を公共部門に応用する格好で、特定のサービスについて様々な評価項目の満足度・重要度を把握するという取組が、米国の1990年代の成功事例を踏まえて²⁸、日本でも普及している面があるが²⁹、国民生活の全般に関わる数十の項目について充足度・満足度を調査してニーズを測るという調査は経済企画庁の国民生活選好度調査として1975年から継続的に行われている（経済企画庁 1976）。

満足度×重要度の分析について、ここでは4点の課題を指摘したい。まず、総合計画の政策や施策のレベルの数十本の項目について満足度・重要度を尋ねる際の項目の表現が簡単に過ぎる場合がある。例えば「学校教育について」「防犯対策について」「工業の振興について」といった紋切り型の表現で、回答者は当該分野の都道府県の役割について、どのようなイメージを描けるだろうか。望ましいのは、政策・施策の効果が住民生活においてどのように発現するのかを示すことであろう。

重要度については、項目リストの中から特に重要なものを選ぶ形式とする場合がある。この方法でも各項目の重要度について一定の情報が得られることになる。さらに、重要性について尋ねずに、生活全体の満足度と項目別の満足度から計量的に、どの項目が全体の満足度をどれだけ左右するかを求め、その影響力の強さを測ることも行われる。ただし、これはいわば潜在的な重要度の把握

であり、項目によっては直接尋ねた重要度と乖離する可能性がある³⁰。また、結果数字を使う側が、計算の意味を理解しておく必要がある³¹。

3点目は、重要度・満足度の2次元の配置から、各項目のニーズや優先度をどのように読むか、である。1枚のグラフを眺めるだけでは心許ない。項目によっては年齢により重要度・満足度が大きく変わる。重要な属性別のプロットは必須であろう。また、一般に収入に関する項目は満足度が低く重要度が高くなる傾向などが早くから指摘されている（例えば経済企画庁 1984）。調査票設計及び結果解釈において、全国と地域、都道府県の政策・施策の役割といった観点が求められる。

4点目は、重要度と満足度のギャップからニーズ度をどのように把握するかという点である。計算するとすれば、重要度－満足度、重要度×不満度（1～5の5段階評価であれば、不満度＝6－満足度）の2通りが典型的な方法である³²。ただし、2次元のプロットにおける各項目の分布に直接対応するのは前者である。後者の計算結果の大小は目で見るイメージとは異なる（図2）。

もう一つ、都道府県で多く行われている住民調査が、CVM（Contingent Valuation Method, 仮想市場法）である。その結果が積極的に公表されることはないのが現状だが、公共事業の事前評価、再評価において、事業の便益を推計する方法として行われる。CVMは、市場のない財・サービスの便益を、あたかも市場があるかのように評価するもので、具体的には、何かの水準を向上させるために支払ってもよいと考える金額（支払意思額）や、水準の悪化をいくらかで受け入れられるか（受

図2 2通りの計算式によるニーズ度の計算例

ニーズ度＝重要度－満足度						ニーズ度＝重要度×不満度 （ただし、不満度＝6－満足度）					
重要度						重要度					
5	4	3	2	1	0	5	25	20	15	10	5
4	3	2	1	0	-1	4	20	16	12	8	4
3	2	1	0	-1	-2	3	15	12	9	6	3
2	1	0	-1	-2	-3	2	10	8	6	4	2
1	0	-1	-2	-3	-4	1	5	4	3	2	1
	1	2	3	4	5		1	2	3	4	5
						満足度					

（出所）筆者作成

入補償額)をアンケートで尋ね、その回答の代表値などに対象人口を乗じて推計する。単純な手法により便益を直接貨幣尺度で捉えるものであり、現在存在しない環境水準の評価など、便益の推計手法の中で最も適用範囲が広い(肥田野 1999)。

適用可能範囲が広いとはいえ、現実の公共事業評価の手法としての課題も多いのであるが³³、ここで指摘しておきたいのは、それ以前の問題である。ワーディングや標本設計がアンケート手法の基本であることはいうまでもないが、自治体がCVMを実施する際には、この基本が危うい状態にある。なぜなら、住民の満足度を計測するための調査とは異なり、調査主体が事業実施主体であり、調査実施の目的が、客観的な推計よりも、国の補助金を受けるのに適切な水準の便益の裏付けや、事業の実施・継続に関して審議する第三者機関を納得させるためのデータを得るということに(真摯に事業を執行したいという姿勢があったとしても、あるいはその姿勢の故に)なりがちだからである。CVMでは個々の事業ごとに、便益計測の対象とする範囲をどのように決め、そこから標本をいかに抽出し、調査票ではどのようなワーディングで事業目的や便益を回答者に説明するのかが、結果数字を大きく左右する。CVMのバイアス要因については整理され、世の中にはガイドラインも提案されているが(肥田野 1999)、そもそも調査を客観的に行おうという意思がなければ、どのような技術的検討も役に立たない。

なお、CVM以外に、公共事業評価において、事業に関するニーズや要望、満足度などをアンケートで計測することも様々行われているが、上と同様のバイアスの危険がある。

この状態の改善に向けては、外部からのチェック機能が有効であると考えられ、総務省行政評価局による点検や全都道府県に設置された第三者機関が果たすべき役割が大きいはずである³⁴。

7. 公共事業の費用便益分析の課題

一般の行政評価・政策評価とは切り離された形が多いながらも、独自の視点で評価する仕組みを設けたり、第三者機関が、義務付けられている再

評価以外の評価にも関わったりする団体が増えていることは既に述べた通りである。

都道府県が独自の視点で、自ら行う公共事業を評価しようとするれば、都道府県にとっての効率と全国の効率とでどう折り合いをつけるか(他県の負担による効果を享受すべきか)、公共事業に使う1円と一般事業に使う1円について、同じだけの吟味がなされているか、など難しい課題も多い。

公共事業の評価手法の中心はいうまでもなく費用便益分析であるが、すべての要素を貨幣価値に換算することの限界もまた明らかである。実務上では一般に、できるだけ多くの要素を貨幣価値化して費用便益分析の精緻化を図るアプローチと、費用便益分析に加えて、貨幣価値化が困難な要素を評価項目として取り込む「多基準分析」のアプローチの2つがあり、多基準分析はさらに非貨幣的項目を点数化するか否か、総合点を算出するか否か(総合点を算出する場合を狭義多基準分析と呼ぶ。またしばしば総合評価方式とも呼ばれる)によって分類することができる(山崎 2006、松田・荻野・塚田 2004)。

都道府県の独自の取り組みには、費用便益分析を独自に精緻化する例もあるが、多基準分析が多い。国土交通省も「公共事業評価の基本的考え方」(公共事業評価システム研究会 2002)において、「公共事業は多様な視点から評価されるべきである」とし、多段階の階層構造をもつ諸要素を総合化して評価する総合評価を提案している。

しかし、この多基準分析、特に総合評価方式については、金本(2003)が指摘するとおり、ダブルカウントが起きやすい、費用便益比が1をはるかに下回る事業でも高得点になりうる、評価項目によっては不確実性が大きく総合点の信頼性を損なう、などの問題が明らかである。

ところで、この多基準分析は、客観的に計測される評価指標と、評価者が主観的に自己評価・判定する項目を組み合わせるという点において、典型的な事務事業評価と同じ方法であるといえる。実際、そのような観点から公共事業と一般事業を共通のフォーマットで評価しようとする取り組みもある。しかし、ここで指摘しておく必要があるのは、一般の事業と公共事業の性格の違いである。一般の事務事業評価における観点別の自己評価を

真に客観的な評価結果と思う人はいないだろうが、それでも評価作業を論理的あるいは誠実に言えばPDCAサイクルのチェックとして機能する。しかし、ひとたび着手すれば後戻りや方向転換が難しく、事業規模も大きい公共事業では、常に厳密な客観性が求められるといえ、事業担当者の自己評価ではこの要請に応えることができない。

したがって、公共事業の評価はできるだけ費用便益分析を中心に行うことが望ましいであろう。だが、都道府県において費用便益分析を行う場合の課題もまた多い。まず、国の補助事業に関して求められる費用便益分析は、省庁別・事業分野別に定められたマニュアル類³⁵にしたがって行われるが、同じマニュアルに基づいて計算された効率を比較すること（省庁における作業）はできても、異なるマニュアルによって算出された効率を比較することができない（大野 2002、2005）。例えばマニュアル通りに計算した一般道路と広域農道の効率は比較することができないのである。

また個々のマニュアルに定められた計算方法も、全国の事業を横断的にできるだけ簡素な方法で比較するという目的のためのものであり、地域の状況によっては非現実的な計算となる場合がある。例えば、最も費用便益計算が明確にできるとされる道路事業について、計算する3種類の利用者便益の1つに交通事故減少便益がある。ところが、この便益は走行台数、延長キロ、交差点数の3要素によって決まる（国土交通省 2003）ため、ある道路の交通事故多発をうけてバイパスを整備する際に、迂回によって延長と交差点数が減らなければ、便益はマイナスになってしまう。また便益計算の前提として必要な交通量は、将来的な道路網における、道路交通センサスの結果に基づく交通量の2030年推計を、全国を数千（おおむね市区町村程度）に分割したゾーン単位に配分したものが基本となるが、道路交通センサスの対象とならない道路（市町村道や農道）の役割が大きいような小さな地域では、現実的なものとならない。

そもそも、事業を担当する省庁・部署が評価のためのマニュアル等を作成するため、必ずバイアスが出てくる（金本 2003）という現実がある。実際、第一次産業の、就業者の高齢化が著しい地域の事業において、年間便益を将来50年間にわた

って不変（現在価値には換算するとして）とするような「推計」が認められている。

では、都道府県としてはどのような評価をすべきか。まず、できるだけ独自の視点で、国のマニュアルをもとにしつつも、費用と便益をできるだけ地域の実情に合うように計算する必要がある。貨幣価値にできないものも、定量的に効果を推計する。あまり表立って議論されない地元建設業への効果（本音の効果として間違いなく大きい）や地元の実質負担額（直接の補助金のほかにも別途の手当がある場合が多い）³⁶についても、できるだけ明確にする必要がある。その上で、やはり事業実施主体が計算することのバイアスもまた避けられないのであり（金本 2003）、外部からのチェックが必須である。CVMと同様、総務省行政評価局の点検や全都道府県に設置された第三者機関が果たすべき役割が大きい³⁷。

8. 定量評価の今後に向けて

総じて、定量評価においては、手法のもつ技術的な限界と、不十分・不適切なデータの取り扱いに起因する問題を峻別することが重要である。後者の問題は、いわば「評価において、統計でウソをつく法」³⁸として、次の4通りに分類できる—①ウソのデータ（非現実的なデータ）をつくる、②データを恣意的に使う、③データを誤って使う（あるいは使い方を誤る）、④（必要な説明を欠くことによって）数字を独り歩きさせる。本稿で取り上げた問題点も、純粋に技術的なものを除いて、すべて①～④のどれかにあてはまる。対策として、①と②は、与えられた職務を良心的に遂行しようとする姿勢が背後にある場合も多く、外部からのチェックこそが有効な対策であろう。③と④は、内部における評価能力の向上がまず求められ、その上で、外部からのチェックが必要となろう。

ここで外部からチェックをするのは、設置が義務付けられている再評価など公共事業評価のための機関のほか、一般の行政評価・政策評価の二次評価を担うなど、すでに多数設置されている第三者機関となるが、その活用は十分とはいえない。また研究者による外部からのチェックも大いに期

待されるが、そのためには評価情報の全面的な公開が前提となる³⁹。

内部の評価能力向上について、定量評価の問題を一朝一夕に解決することは期待できないが、ごく基本的な事項を確認するだけで、事態は大きく改善するはずである。指標やアンケートの妥当性・信頼性に関するチェックリストや、データ本体と補助情報を必ず一体のものとする仕組みなどが有効であろう。また評価活動の支援やデータ収集に携わる調査会社やコンサルタントには、自ら統計でウソをつくようなことがないようにしてもらわねばならない。

質の低い定量評価は、その結果を活用しようとすれば直接意思決定者の誤った判断を導く危険があり、その数字が使えないと実務者や市民によってみなされてしまえば、評価システム自体の信頼を失う結果を招くことになりかねない。

注記

- 1 行政評価・政策評価の定着に時間を要することは当初から指摘されていたが、一定の試行期間を経て本格導入に至って後も、様々な改定を繰り返す団体が多い。評価対象を増す、フォーマットを改定する、評価作業を簡素化する、総合計画と一緒に評価制度も切り替えるなど、改定の中身は多岐にわたる。
- 2 現在では、ほとんどの都道府県が行政評価・政策評価の情報を積極的に公表しており、HPも評価結果をすべて掲載するなど、広く活用されているが、次の団体については詳細を把握することができなかった一山形県（ごく断片的な情報のみ掲載）、新潟県（評価制度を見直し中で、今年度からあらためて実施予定。過去の評価情報も見つからず。ただし総合計画の進行管理に関する情報はあり）、岐阜県（事務事業評価を見直し中で、今年度から再開される模様。過去の評価情報も見つからず）、大阪府（「施策評価の結果（詳細）」については、現在メンテナンス中」とのことで、事務事業の評価結果、過去の評価結果を含め詳細わからず）、兵庫県（行政評価を見直し中と思われ、該当ページには「現在、この分野の情報はありません。」と表示。過去の評価情報もなし。総合計画の「社会像評価」については、「現在、ホームページ移行作業中」のため見られず）、

高知県（評価を行っているか否かを含め、情報得られず）。

なお、上記の団体を含め、全部の団体について個別に詳細を問い合わせることが難しいため、今回の調査はHPから得られる情報のみに基づいている。

また、今回の調査の過程で、上記で言及した団体を含め、評価制度・システムを変えた段階で、以前の制度・システムの評価結果や評価方法に関する情報を消してしまう団体が少なくないことがわかった。

- 3 特定の事業や個別分野の計画でなく、分野を横断して体系的・統一的に評価する仕組みを行政評価・政策評価とした。したがって、総合計画の進行管理など、評価と呼ばれていなくても、その方法によっては評価制度とみなしている。また公共事業を一括して評価対象とする制度も本稿では行政評価・政策評価とした。
- 4 その他に用いられている総称としては、政策評価以外の「政策〇〇」というタイプが3団体、施策評価・事務事業評価など「施策」「事業」「事務事業」という語を用いるタイプ（評価対象のレベルと合致するわけではない）が7団体、など。
- 5 注3で述べたように、総合計画の体系的・統一的な進行管理の仕組みを行政評価・政策評価とみなす本稿の観点からは、行政評価・政策評価の実績について、さらに遡ることができる。例えば、厚木市は1989年の時点ですでに施策効果を評価指標で把握しようとする試みを行っている（斎藤1994）。
- 6 総務省の調査では、行政評価を「政策、施策、事務事業について、事前、事中、事後を問わず、一定の基準、指標をもって、妥当性、達成度や成果を判定するもの」と定義している。ただし、旧自治省時代の調査では、「国が定めた評価制度は除く」とされていた。総合計画の進行管理を行う評価システムなどは、団体内で評価制度と呼ばれていない場合、総務省の調査で評価制度の1つとして扱われていない場合があると考えられる。また公共事業評価も、独自の評価システムをもっている場合を含めて、調査結果に含まれていない場合があると考えられる。
- 7 総務省の調査において、現在まで一貫して行政評価を導入していないのが鳥取県である。片山前知事は2期の在任中、他の団体のような行政評価制度を導入しないことを明言していた。また、近年では多く

の団体が行政評価・政策評価の方式で総合計画の進行管理を行っているが、鳥取県では、片山前知事の在任中、総合計画を策定していない。ただし、複数の個別分野の計画で目標値設定と達成度評価を行うことになっており、また公共事業の外部評価には、本文で後述の通り、かなり徹底して取り組んでいる。

なお、2006年度は導入済みの団体が46から45に減少しているが、これは注2でも触れた新潟県が、評価システムを見直し中であることによる。

- 8 階層のどのレベルを評価対象としているか、そもそも政策・施策・事務事業の区分に統一的な理解があるわけでもなく、自治体の分類もまちまちである。また、政策・施策レベルの評価といっても、その内容は、一定の掘り下げた検討を行っているものから、下位階層の施策・事業の評価結果の抜粋を一覧にまとめただけのものや、評価指標のモニタリングをしているだけのものまで様々である。今回のHPを通じた調査でも、総務省の調査結果が実態と合わないケースがいくつか見られたが、ここで述べている大勢に影響はないと考えられる。
- 9 総務省の06年10月時点の調査結果の数を大きく上回っているのは、今年度新しい基本計画等のスタートにあわせて評価制度を総合計画に対応するように改定した団体、総合計画の進行管理用の指標を設けていてもそれが総務省の調査で評価としてカウントされていない団体がそれぞれ複数あることによる。
- 10 残りの4団体は、タイプがどちらでもない、あるいは不明である。
- 11 一般に費用便益分析とは貨幣価値に換算できる効果と費用を比較するもの、費用対効果分析とは費用便益分析に加えて貨幣換算できない効果や定性的な効果も加えて分析するものとされるが、中央省庁の実務において両者は区分しないで使用されている（総務省行政監察局 2000）のが実情である。本稿では、主たる関心が本来の費用便益分析にあるため、本文の以下では費用便益分析という語を用いる。
- 12 1997年1月に導入された、北海道の時のアセスメントが、公共事業評価の進展に役買ったことは広く知られている。
- 13 公共事業評価については、HP上に詳しい情報を掲載している団体は多くない。事後評価については、9団体の他に、12団体において事後評価を行っているとの記述があったが、その評価活動の実態が不明

であればカウントしておらず、第三者機関に事業の概要を説明して質疑を行うだけの場合、本稿では評価制度とは呼ばない。なお、事前評価・再評価についても同様の扱いをしており、「評価を行っている」という記述が見られるが表4に含めていない団体が事前評価で5、再評価で12ある。

- 14 「めざそう値」など行政が設定する通常の目標値と区別する意味を込める呼称もあるが、実績値と比べてベンチマーキングすること自体は変わらない。
- 15 指標を重視する程度、また目標値を設定している指標の割合にかかわらず数えている。
- 16 呼称は到達度、進捗率など様々である。複数の「達成度」を計算する団体では別の呼称によって区別している。
- 17 表7に挙げた団体数は、各取り扱いが多少とも見られればカウントしたものであり、質と量においてどの程度徹底しているかを斟酌していない。
- 18 業績測定の最も有名な事例の1つであるオレゴン・ベンチマークスでも、過去のデータ等に基づいて目標値のベースラインを決め、関係部局間の交渉によって最終的に決定するとされている。しかし、ここでも、オレゴン・ベンチマークスの運用を担当するオレゴン・プログレス・ボードの事務局長（当時）のジェフリー・トライエンス氏は2000年の来日時の講演の際、筆者の質問に対して、各部局は自らの仕事との関係が直接的な指標については目標を常に低く設定しようとし、関係があまり直接的でない指標では目標を高めにしようとするとして述べている。
- 19 指標の妥当性の記述がある団体と同数であるが、両条件を満たすのは15である。その中では鹿児島県の施策評価票が、成果と指標との関連、目標値設定の根拠の両方について比較的丁寧な記述をしている。
- 20 斎藤達三は早くからフロー指標とストック指標の違いに言及している。例えば、斎藤（1994）ではストック指標について、値が減少する可能性がないので「その分より厳しい評価基準が適用されなければならない」と指摘している。
- 21 達成度の値によって評定・集計しなくても、達成状況を直接比較すれば、そのことが問題となる。
- 22 分類した後の取り扱いがより重要である。岡山県では、本文⑤で後述の、ストック指標をフロー指標に変換する方法によりうまく処理している。
- 23 目標達成度の計算方法の選択に言及している文献は

少ないが、斎藤編（1999）は「目標達成率」を本稿で言う期間型で定義し、期末型の計算を「実現した現時点の状態を示す指標」に過ぎず、成果の評価尺度として用いるのは「よく生じる初歩的な誤解」であると主張している。また、関田・加藤（2001）は期末型の達成度を「達成度指数」、前年度からの1年間を対象とする期間型の達成度を「限界的達成度指数」と呼び、両者を組み合わせる評価方法に言及している。

- 24 2種類の計算式の使い分け方は、施策レベルを期間型、事務事業レベルを期末型とする（山梨県、愛媛県）、目標達成までは期間型、目標達成後は期末型（長崎県）など。全プログラムについて2通りの計算を行っているのが岡山県の「夢づくり政策評価」で、期末型の「達成率」（注22で述べたように、ストック指標は1年間の期間型を計算しており、フロー指標に変換した期末型と同値）と期間型の「進捗率」をすべての「快適生活指標」について計算しており、フロー・ストックの区別を含め、最も無難な取り扱いをしているといえる。
- 25 厳密にはすべてが順序尺度で、間隔尺度の評定は一つもない。ただし、達成度「100%以上」と「25%以下」の間を等間隔に区分するなど、間隔尺度に近い評定をしている団体も少数ある。
- 26 宮城県の2007年度の県民満足度調査は、満足度を4段階評価に変えるなどの変更が行われている。
- 27 マーケティングの分野では古くから、顧客満足はサービスの期待水準と実際に認知した水準のギャップによって決まるという議論がなされている（Parasuraman et al. 1985, Maister 1985）。また、経済学の分野でも、主観的厚生指標について、ある時点の集団内では所得が高いほど指標も高いが、時系列では所得が向上しても指標はほとんど上昇しないという実証的な報告がある（Frey and Stutzer 2002, Layard 2003）。
- 28 上山（1999）が、米国の成功事例を紹介している。
- 29 初期の取組として、2000年の逗子市（逗子市 2000）や2001年の杉並区（谷口 2001）などを挙げることができる。
- 30 梅田・小野・中泉（2004）に具体例がある。
- 31 この種の計算はしばしば回帰分析（を調査会社等が行うこと）によってなされる。計算結果を活用するためには、関係する職員が回帰分析の意味や結果数

字の制約をある程度理解している必要がある。

- 32 ニーズの把握に工夫を凝らす場合、結果が使えるものでなければ意味がない。某県では、重要度と実現度からニーズ（優先的に実現に努める必要性）を把握する方法として、実現度の上げ易さを考慮すると、疑問符のつく大胆な仮定をいくつか置いた上で、重要度×（5－実現度）^{1/2}×役割度という計算を行っているが、趣旨からいえば重要度×（5－実現度）^{3/2}×役割度とするのが自然である。
- 33 国土交通省に2007年1月設置された「公共事業評価手法に関する検討会」でも議論されている。
- 34 総務省行政評価局は、2006年度の客観性担保評価（政策評価の点検）において、国土交通省所管の事業のCVMによる便益の過大推計を指摘しているが、1件の指摘に留まっている（総務省行政評価局 2007）。また筆者が現在会長を務める鳥取県公共事業評価委員会は、ある事業の便益推計に用いられたCVMの過大推計を指摘し、CVMの実施時に測定の妥当性を失わないように求めている（鳥取県公共事業評価委員会 2006a）。しかしこのようなチェック機能を發揮している第三者機関がどれだけあるか、あまり期待はできないのが現状である。
- 35 総務省（2003）によれば、03年時点で、評価実施要領が57、費用対効果分析マニュアル等が39ある。
- 36 かつては、いかに効率の悪い事業でも地元にとってはお買い得、というモラルハザードの危険が多分にあったといえるが、昨今の自治体の財政事情は、たとえお買い得でも新規事業を賄う財源がない、という状況になりつつある。評価によって真に必要な事業を明らかにする必要が高まったといえる。
- 37 総務省行政評価局は、2006年度の客観性担保評価において、約10件の公共事業評価における便益の過大推計を指摘しているが（総務省行政評価局 2007）、これは氷山の一角であろう。また、鳥取県公共事業評価委員会は、06年度の再評価において、残事業の便益を精査した末、1事業の中止を答申し（県側が想定していない中止を答申した全国的にもまれなケースと思われる、鳥取県公共事業評価委員会 2006a）、事前評価においては、マニュアル通りの計画交通量が現実から乖離しているとして独自の調査を求めた（その後実際に起終点調査（OD調査）等が緻密に行われたが、これもまた全国的にもまれなケースと思われる、鳥取県公共事業評価委員会

2006b)。ただし、このような外部評価を行うためには、十分な回数・時間の審議（年間2、3度の開催で公共事業の外部評価を行うことは不可能である。もっとも鳥取県ではこの年、評価対象事業が15と少なかったため、1件1件を十分に検討できたという事情もある）と事業所管部局の真摯な協力が前提となる。

38 この表現はハフの名著（ハフ 1968）の邦訳書名にならなかった。

39 公共事業評価に関する情報公開はとりわけ不十分である。積極的に評価情報の公開を進めている国土交通省も、第三者が検証するという観点からは不十分な状況である（山田 2006）。今回の調査でも、公共事業評価についての詳細がHP上で公表されている都道府県は少数であり、第三者が検証するのに十分な情報が公開されている団体はほとんどなかった。

参考文献

- 上田尚一（2003）『統計の誤用・活用』、朝倉書店
- 上山信一（1999）『「行政経営」の時代』、NTT出版
- 梅田次郎（2004）「経営改革を目指した行政評価・三重方式と評価指標」、『月刊自治研』、第46巻9月号
- 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
- 大野泰資（2002）「類似公共事業の評価手法の相違がもたらす帰結について」、『会計検査研究』、No.25
- 大野泰資（2005）「公共事業評価の実際―省庁別・事業別の費用便益分析がもたらす結果について」、『フィナンシャル・レビュー』、2005年7月号
- 小野達也（2001）「地方自治体の行政評価システムの課題と成功への条件」、『富士通総研研究レポート』、No.117
- 小野達也（2004）「数字とデータは評価の客観性を高めているか―統計でウソをつけば、評価は偽物となる」、『日本評価学会春季第1回全国大会報告論文集』
- 小野達也（2006）「都道府県の行政評価、この10年間の問題点～定量評価の陥穽～」、『日本評価学会第7回全国大会報告論文集』
- 小野達也（2007）「自治体評価における計量・計数の現状と課題」、『日本評価学会第8回全国大会報告論文集』
- 金本良嗣（2003）「日本における公共事業評価の現状と課題」、『行政管理研究』、No.103
- 経済企画庁（1976）『国民生活選好度調査第2回結果報告書』、日興リサーチセンター
- 経済企画庁（1984）『第4回国民生活選好度調査』
- 公共事業評価システム研究会（2002）『公共事業評価の基本的考え方』
- 国土交通省（2003）『費用便益分析マニュアル』
- 斎藤達三（1994）『総合計画の管理と評価』、勁草書房
- 斎藤達三編（1999）『実践自治体政策評価』、ぎょうせい
- 佐藤徹（2004）「公共サービスにおける評価指標とは何か」、『月刊自治研』、第46巻9月号
- 自治省（2000）『地方公共団体における行政評価の取組状況（平成11年9月末現在）』
- 自治省（2001）『地方公共団体における行政評価の取組状況（平成12年8月末現在）』
- 白川一郎・小野達也（2002）『NPMの導入による地方自治体の改革に向けて』、富士通総研研究レポート、No.140
- 逗子市（2000）『逗子海岸海水浴場におけるCS（顧客満足度）調査報告書』
- 関田康慶・加藤由美（2001）「政策・施策評価システムの設計と評価方法」、『会計検査研究』、No.24
- 総務省行政監察局（2000）『公共事業の評価に関する調査結果報告書』
- 総務省行政評価局（2003）『公共事業に関する評価実施要領・費用対効果分析マニュアル等の策定状況』
- 総務省行政評価局（2007）『政策評価の点検結果（平成18年度）』
- 総務省自治行政局（2001）『地方公共団体における行政評価の取組状況（平成13年7月末現在）』
- 総務省自治行政局（2002a）『行政評価指標設定の課題と考え方』
- 総務省自治行政局（2002b）『地方公共団体における行政評価の取組状況（平成14年7月末現在）』
- 総務省自治行政局（2003）『地方公共団体における行政評価の取組状況（平成15年7月末現在）』
- 総務省自治行政局（2004）『地方公共団体における行政評価の取組状況（平成16年7月末現在）』
- 総務省自治行政局（2006）『地方公共団体における行政評価の取組状況（平成18年1月1日現在）』
- 総務省自治行政局（2007）『地方公共団体における行政評価の取組状況（平成18年10月1日現在）』
- 田辺国昭（2001）「実績評価の課題」、『都市問題研究』、

53 (10)

- 谷口俊治 (2001) 「より良い行政サービスのための「行政CS評価システム」」、『地域経営ニュースレター』
野村総合研究所、33
- 鳥取県公共事業評価委員会 (2006a) 『平成18年度公共事業の再評価について (答申)』
- 鳥取県公共事業評価委員会 (2006b) 『平成18年度公共事業の事前評価について (中間答申)』
- ハフ, D. (1968) 『統計でウソをつく法』、講談社
- 肥田野登 (1999) 『環境と行政の経済評価—CVMマニュアル』、勁草書房
- 松田和香・荻野宏之・塚田幸広 (2004) 「諸外国および地方公共団体における総合的な道路事業評価手法の特徴」、『土木計画学研究・講演集』、30 (131)
- 山崎治 (2006) 「公共事業の事前評価」、『レファレンス』、(661)
- 山田宏 (2006) 「公共事業における費用便益分析の役割」、『立法と調査』、(256)
- 吉川富夫 (2004) 『米国における地域経営の新展開—業績測定による成長管理』、公人社
- Affholter, D.P. (1994) . Outcome Monitoring. In J.S.Wholesy et al. (Eds) , *Handbook of Practical Program Evaluation*. San Francisco, Calif: Jossey-Bass.

- Frey, B., and Stutzer, A. (2002) . What Can Economists Learn from Happiness Research?. *Journal of Economic Literature*, 40, 402-435.
- Hatry, H. P. (1999) . *Performance Measurement: Getting Results*. Washington, D.C.: Urban Institute.
- Layard, R. (2003) . Happiness: Has Social Science a Clue?. A paper presented at *Lionel Robbins Memorial Lectures* on 3,4,5 March 2003, London School of Economics.
- Maister, D.H. (1985) . The Psychology of Waiting Lines. In J.A.Czepiel, M.R.Solomon, and C.F. Suprenant (Eds) , *The Service Encounter*. Lanham, MD: Lexington Books.
- Parasuraman, A., Zeithaml, V.A. and Berry, L.L. (1985) . A Conceptual Model of Service Quality and Its Implications for Future Research. *Journal of Marketing*, 49 (4) , 41-50.
- Wholey, J.S. (1999) . Quality Control: Assessing the Accuracy and Usefulness of Performance Measurement Systems. In H. P. Hatry, *Performance Measurement: Getting Results*. Washington, D.C.: Urban Institute.

(2008.2.1受理)

Ten Years of Evaluation Practices in the Prefectures of Japan –The Past and Future of Quantitative Methods–

Tatsuya Ono

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

Abstract

Prefectures have played a leading role in the introduction of evaluation systems into local governments in Japan. At present, almost all local governments have a variety of evaluation systems, most of which adopt quantitative methods to obtain objectivity in the evaluation results. However, there seems to be a number of problems in their handling of data, performance measures and quantitative analysis. This article presents the results from a survey of evaluation systems and quantitative methods used in all the prefectures of Japan. Typical problems are found in the target setting and benchmarking in performance measurement, resident surveys including customer satisfaction surveys and contingent valuation methods, and cost-benefit analyses of public works. This article also discusses how to cope with these problems in evaluation practice.

Keywords

prefectures, quantitative methods, performance measurement,
resident survey, cost-benefit analysis

【研究論文：依頼原稿】

都市自治体の評価：本格普及から10年後の実態

田中 啓

静岡文化芸術大学

hiraki@suac.ac.jp

要 約

日本の自治体評価は、その本格的な普及が始まってから10年余りが経過した。本稿は、その自治体評価の実態を把握することを目的とする探索的研究（exploratory study）である。本稿では、筆者が実施した自治体評価に関する調査結果を利用して、都市自治体（市・特別区）を中心に行政評価の実態把握を試みた。特に、行政評価の実施による成果や行政評価の抱える問題点の把握に主眼を置いた。その結果、現在の自治体評価は職員の意識変化や事務事業レベルの改善面では有効性を発揮しているものの、全庁的な資源配分の改善面ではあまり有効ではないこと、評価の質や職員の評価能力について問題を抱えている都市自治体が少なくないことなどが示唆された。本稿の分析を踏まえると、今後の自治体評価については「脱事務事業評価」「予算と評価の連動」「脱業績測定」という方向性が示唆される。このような方向性が実現するかどうかは、自治体自身の認識や取り組みのあり方のほか、評価研究者の貢献にかかっている。

キーワード

自治体評価、業績測定、行政評価の成果、評価と予算の連動

1. 転換期にある自治体評価¹

国内に「行政評価」という仕組みが本格的に普及しはじめてから10年余りが経過した。当初は都道府県や政令指定都市をはじめとする比較的規模の大きい自治体において行政評価への取り組みが目立つ程度であった。しかし、2000年を超える頃から行政評価への関心が自治体の間に急速に広まっていき、小規模な自治体においても行政評価制度が導入されるようになった。この結果、総務省の調査によれば、2006年10月時点で、ほとんど全て（96%、45団体）の都道府県、約7割（67%、

427団体）の市・特別区、3割近く（27%、169団体）の町・村がそれぞれ行政評価制度を導入するに至っている（総務省 2007）。

自治体に行政評価が普及しはじめてから10年余りが経過した今日、自治体の行政評価が転換点を迎えているのではないかとというのが本稿の基本認識である。1990年代後半（1996～99年頃）を自治体における評価の「導入初期（揺籃期）」、2000年代前半（2000～04年頃）を「本格導入期（ブーム期）」とすれば、そろそろ新しい局面を迎えつつあるのではないかと考えられるのである（表1を参照）。その主な理由は、以下の3点である。

表1 自治体における行政評価の展開プロセス

行政評価のフェーズ	1990年代後半 (1995～99年頃)	2000年代前半 (2000～04年頃)	2000年代後半～ (2005年頃以降)
	導入初期（揺籃期）	本格導入期（ブーム期）	本格運用期（多様化期）
主要な課題	・評価の概念・手法等の認知・普及 ・日本に適した評価手法の模索	・行政評価の定型化・標準化 ・行政評価の運用効率化	・行政評価の活用 ・行政評価の質の確保
主な手法	・事務事業評価	・事務事業評価が中心 ・政策・施策評価の登場 ・新しい手法の登場（戦略計画、BSC、ベンチマーキング等）	事務事業評価を中心としながら、有効な手法や活用方法を模索（現時点では不透明）
関連する動き	・三重県が事務事業評価システムを導入（1996年） ・行政改革会議最終報告（1997年）	・国の政策評価制度開始（2001年） ・独立行政法人設置（2001年） ・政策評価法施行（2002年） ・市町村合併の推進（1999年以降） ・「今後の行政改革の方針」（2004年）	・政策評価制度の見直し（2005年度） ・新行革大綱と集中改革プラン（共に2005年度）

（出所）田中（2005b）

第1に、自治体における行政評価の導入がそろそろ一段落しつつあるということである。平成の大合併の影響により、導入時期が後ろ倒しになった自治体があると思われる一方で、普及率の低かった町村が合併により減少したことを考えれば、むしろ全体としてみれば、行政評価の普及度合いが一気に高まったと言える。

第2に、自治体における行政評価の普及率の高まりや実施経験の蓄積を反映して、評価への取り組みの内容や質に変化が生じている。先行した自治体を中心に、既存の評価制度を改変・拡充する取り組みが広がる一方、戦略計画、バランス・スコアカード、ベンチマーキングなど新しい取り組みも始まっている。また、先行自治体における成功・失敗経験とそれらに基づく評価のノウハウは、さまざまな形で自治体の間で共有され、自治体全体の評価に関する知識・ノウハウの底上げにつながっている。

最後に、地方分権改革が進展する中で、自治体が行行政評価に対してより積極的に取り組むべき誘因が存在する。合併を選択したにせよしなかったにせよ、今後も多くの自治体にとっては財政面を中心として厳しい環境が続くことが予想される。こうした中で、各自治体における自律的な改革への取り組みが必要になっていることに加えて、自治体が行行政評価を導入することが国の方針として

明示されている（総務省 2005）こともあり、行政評価を導入・活用して改革を進めていこうとする自治体が増加することが予想される。

以上を総合すれば、自治体の行政評価は、その導入・普及が課題であったこれまでのフェーズから、「本格運用」という新しい段階に移りつつあると考えられる。言い換えれば、これまでは多くの自治体が行行政評価を導入することをめざしてきたが、これからは、導入した行政評価制度を活用して、実際に成果を上げていくべき段階に移りつつあると考えられるのである。このような問題意識の下で、本稿の研究は構想された。

2. 本稿の目的

自治体に評価が本格的に普及しはじめてからおよそ10年が経過した節目の時期に、自治体評価の軌跡を振り返ること自体、有意義なことであろう。さらに、もしも本稿の認識のとおり、自治体評価が転換期を迎えているとすれば、過去の実施状況を分析することにより、自治体評価が今後めざすべき方向性について、有益な示唆を得られる可能性がある。

以上の認識に基づき、地方自治体における行政評価の実態を把握することが本稿の目的である。

それは、行政評価の現状を多様な側面から分析することを意味するが、本稿では特に、自治体が行政評価を実施したことによる成果や行政評価の抱える問題点に注目する。したがって、本稿は特定の仮説を検証することを目指すものではなく、むしろ自治体の行政評価に関する新たな仮説の構築につながるような知見を得ることを目的としている。その意味で、本稿は探索的研究（exploratory study）である。

本稿の目的を達成するために用いたのは、筆者が独自に実施した調査の結果である。自治体評価に関する類似の調査は、先に言及した総務省によるものも含めて数多く実施されている²が、筆者の実施した調査が既存の調査と異なるのは、行政評価を実施した成果や行政評価制度の問題点について、詳しく把握を試みている点である。

なお本稿では、地方自治体のうち特に都市自治体（市と特別区）に注目し、これらの自治体における行政評価の実施状況とその成果や問題点の把握を行い、自治体評価の今後の方向性について示唆を得ることとした。

都市自治体においては、行政評価が既に広く普及している上に、今後も（少なくとも数の上では）都市自治体が行政評価の主要な実施主体であり続けると考えられる。また都市自治体といっても、その規模には大きな幅があり、行政評価への取り組み期間や取り組み方法にも多様性がみられる。都市自治体にみられるこうした多様性が、自治体評価の今後の方向性について豊富な示唆を提供してくれるものと期待できる。

3. 都市自治体における行政評価の実施状況

(1) 「実態調査」の概要

前述したように、自治体評価の実態を把握するために、国内の都道府県・市・特別区を対象として、「行政評価の実態に関する調査」（以下「実態調査」）という独自の調査を2006年3月に実施した。この「実態調査」においては、既存の調査と同様、自治体における行政評価の導入状況や制度的特徴

を把握するほか、行政評価を導入した成果や改善すべき問題の所在についても把握を試みた。

なお、この「実態調査」では、アンケート形式の調査票を自治体に直接郵送し、各自治体において回答後に返送してもらうという方法（郵送調査法）を採用した。各自治体においては、「行政評価制度の導入・運用面で主導的な役割を果たされている方（推進部門の長や担当者の方など）」（調査票に記載した表現）に回答してもらうよう依頼した。

したがって、以下に示す調査結果は、あくまで各自治体における行政評価の推進担当者の認識を反映したものであり、それが各自治体の行政評価の実態を正確に反映しているとは限らない点には留意が必要である（読者はこの点に留意して本稿を読んで頂きたい）³。

とはいうものの、各自治体における行政評価の推進担当者は、自団体の評価について最も熟知している職員と目されることから、「行政評価の推進担当者がどのように認識しているか」を把握することは、自治体評価の実態に近似的に接近する手段としては、妥当なものだと考える（なお「実態調査」の具体的な実施方法や回収状況については、本稿末の参考資料の付表1、2を参照されたい）。

「実態調査」に対しては、全国の都道府県、市、特別区の68%（564団体）から回答を得ることができ（付表2）、そのうち285団体（35都道府県、250市・特別区）が既に行政評価を「導入済み」であった（表2）。

本稿の関心の対象は都市自治体であることから、以下では、行政評価を既に導入済みであると回答した250の市・特別区に注目し、これらの団体について主要な調査結果とその解釈等を示す。また対比のために、やはり行政評価を導入済みと回答した35の都道府県の調査結果も併せてみることにする。

(2) 行政評価の導入目的と実施期間

行政評価の導入目的は、自治体が行政評価に取り組む理念や方針を示すものであり、かつ導入する行政評価制度の制度設計にも関わることから、極めて重要な観点である。また行政評価の成果を

表2 回答自治体における行政評価の導入状況

	市・特別区		都道府県		全 体	
	回答	割合	回答	割合	回答	割合
導入済み又は試行中	321	60.8%	35	97.2%	356	63.1%
導入済み	250	47.3%	35	97.2%	285	50.5%
試行中	71	13.4%	0	0%	71	12.6%
未導入であるが、導入予定あり	68	12.9%	0	0%	68	12.1%
未導入であるが、導入を検討中	122	23.1%	0	0%	122	21.6%
未導入であり、特に検討も行っていない	13	2.5%	1	2.8%	14	2.5%
その他	4	0.8%	0	0%	4	0.7%
合 計	528	100%	36	100%	564	100%

(出所) 田中 (2006a)

検討する際には、自治体が行政評価を導入した目的と対比する必要がある。

表3には、導入目的別の回答自治体数と割合が示されている（60%を超える自治体を選択した項目は網掛けで示した）。これによれば、市・区においては、導入目的として「事務事業の廃止・削減」「事務事業の効率性向上」「事務事業の有効性向上」といった事務事業に関わる目的と、「住民等の行政外の主体に対して行政の活動状況を明らかにする」

かにする」といったアカウンタビリティに関わる目的を選んだ自治体が大半を占めた。

この傾向は、都道府県においてもほぼ同様であった。ただし、都道府県においては「長期計画（総合計画など）の進行管理」を目的として挙げた自治体が市・区に比べて15ポイント程度多かった。

なお1つの自治体を選択した目的の数の平均は4.4個であり、回答結果から判断すると、大半の

表3 行政評価の導入目的

導 入 目 的	市・特別区		都道府県		全 体	
	回答	割合	回答	割合	回答	割合
事務事業の廃止・削減	156	62.4%	21	60%	177	62.1%
事務事業の効率性向上	207	82.8%	26	74.3%	233	81.8%
事務事業の有効性向上	207	82.8%	29	82.9%	236	82.8%
全庁レベルの予算配分の変更	62	24.8%	8	22.9%	70	24.6%
全庁レベルの人員配分の変更	23	9.2%	3	8.6%	26	9.1%
長期計画（総合計画など）の進行管理	118	47.2%	22	62.9%	140	49.1%
官民の役割分担の見直し	79	31.6%	11	31.4%	90	31.6%
住民等の行政外の主体に対して行政の活動状況を明らかにする	180	72%	31	88.6%	211	74%
その他	59	23.6%	7	20%	66	23.2%
合 計	250		35		285	

(注1) 行政評価制度を「導入済み」と回答した自治体のみについて集計。

(注2) 網掛けした項目は、回答した自治体の割合が60%以上のもの。

(注3) 複数回答であるため、各項目（列）の回答割合の合計は100%にはならない。

(出所) 田中 (2006a)

自治体が、事務事業関連の3つの目的とアカウンタビリティに関わる目的（1つ）を行政評価の目的として同時に設定している傾向が明確に表れている。

行政評価を導入する自治体の多くは事務事業評価を採用していることから、大半の自治体が導入目的として事務事業関連の目的を設定しているという調査結果は、事務事業評価が支配的である事実とも整合している。それに加えて、事務事業評価を実施することにより、住民等に対するアカウンタビリティも果たすという「一石二鳥」を多くの自治体が狙っていることも、この調査結果からうかがうことができる。

次に、都市自治体における行政評価の実施期間（行政評価制度を導入後の経過期間）は、4年未満の団体が7割近く（66%）を占めている。行政評価制度を6年以上運用している市・区は16団体（6.4%）に留まっており、8年以上の実施経験のある団体は回答団体の中には存在しなかった（図1）。

全体的には都道府県の方が実施期間が長く、導入後4年以上を経過した団体が7割以上（71%）、実施期間が8年を超える都道府県も5団体（14.3%）存在した。回答団体中、運用経験の最も長かった都道府県（無記名調査のため自治体名は不明）の

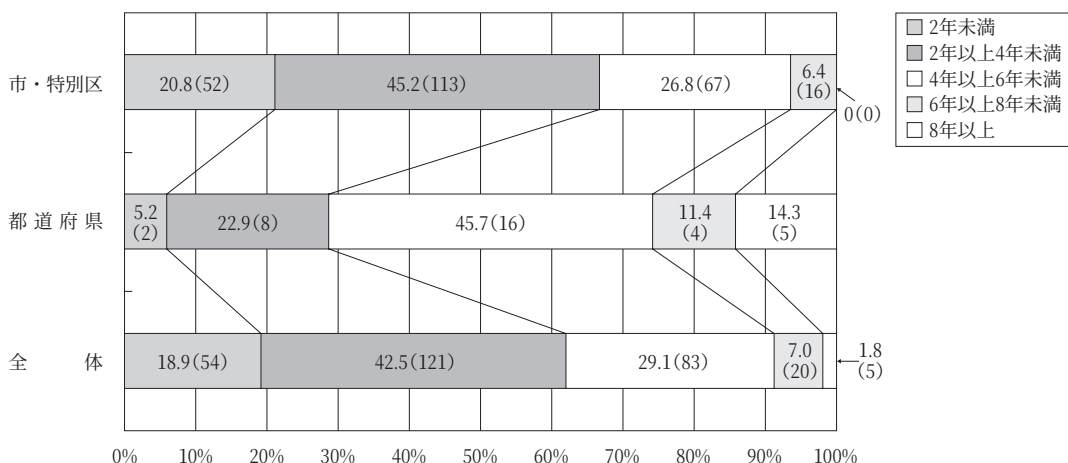
実施期間は9年11か月（2006年2月末時点）であった。また実施期間の平均は、都道府県が5.0年に対して市・区は3.2年と2年近く短かった。これらの結果から明らかなように、全体としては都道府県が先行して行政評価への取り組みを開始しており、8年を超える長期間の経験を蓄積する団体も現れている。一方、都市自治体は団体数が多いこともあり、実施期間の長さは都道府県以上にばらつきが顕著である。

（3）行政評価の質の確保

行政評価の手法の中でも業績測定（performance measurement）においては、設定する評価指標や目標値の質が、評価全体の質や有効性を規定する。古川・北大路（2004）などが指摘するように、その位置づけは自治体によって異なるものの、業績測定を行政評価制度に導入している自治体が多いことから、各自治体が行政評価の質を確保するために、どのような方法で評価指標や目標値の妥当性を判断しているかは重要な観点である。

表4は、自治体がどのような方法で評価指標や目標値の妥当性を判断しているかを把握した結果である。評価指標や目標値を設定するのが1次評

図1 行政評価の実施期間（行政評価制度を導入後の経過期間）



（注1）行政評価制度を「導入済み」と回答した自治体のみについて集計。

（注2）（ ）内は回答数。市・特別区は無回答が2件あったので、各区分の合計は100%にならない。

（出所）田中（2006a）

表4 評価指標や目標値の妥当性判断の方法

判断する主体等	市・特別区		都道府県		全 体	
	回答	割合	回答	割合	回答	割合
庁内の推進部門が行う	114	45.6%	21	60%	135	47.4%
庁内の特別組織(タスク・フォース等)が行う	36	14.4%	3	8.6%	39	13.7%
庁内の2次・3次評価者が行う	72	28.8%	11	31.4%	83	29.1%
庁外のメンバーを主体とする専門組織(第三者委員会等)が行う	26	10.4%	11	31.4%	37	13.0%
明示的に行わない	92	36.8%	10	28.6%	102	35.8%
その他	3	1.2%	4	11.4%	7	2.5%
無回答	1	0.4%	0	0%	1	0.4%
合 計	250		35		285	

(注1) 行政評価制度を「導入済み」と回答した自治体のみにについて集計。

(注2) 複数回答のため各項目(列)の回答割合の合計は100%にはならない。

(出所) 田中(2006a)

価者(多くの場合は評価対象の事務事業等の担当者)だとすれば、1次評価者以外の主体が、設定された評価指標や目標値の妥当性をチェックしているかが焦点となる。

この点については、市・区、都道府県ともに庁内の推進部門が評価指標や目標値の妥当性を判断するとした団体が最も多かった(それぞれ45.6%、60%)。庁内の2次・3次評価者が判断すると回答した団体がこれに次いで多く、市・区、都道府県ともに3割前後(それぞれ28.8%、31.4%)を占めた。都道府県においては、庁外の専門組織にこれらの妥当性の判断を委ねるとした団体も3割を超えた(31.4%)。

一方、評価指標や目標値の妥当性の判断を明示的に行わないとする自治体が少なからず存在し、特に市・区においては4割近く(36.8%)に達した。これらの団体においては、評価指標や目標値の妥当性は1次評価者の判断のみに委ねられていることになり、評価の質を確保する上では問題なしとしない。

評価の質を確保するためには、評価指標や目標値の妥当性だけでなく、評価結果の妥当性についても事後的なチェックを行うことが望ましい。ここで「評価結果」とは、分析結果とそれを踏まえて示された判断・提案の全てを含む。表4と同様の選択肢を用いて、評価結果の妥当性の判断方法

について把握した結果が表5である。

これによれば、市・区は庁内の2次・3次評価者(一般には1次評価者の上長にあたる課長や部長)が妥当性の判断を行うとした団体が約半数(49.2%)を占めた。次いで多かったのは、庁内の推進部門が判断を行うと回答した団体であった(32.4%)。一方、評価結果の妥当性の判断を明示的に行っていない市・区は約4分の1(24.8%)であり、先にみた評価指標や目標値の場合よりは少なかった。都道府県では、庁内の推進部門が判断するとした団体が最も多く(57.1%)、庁外の専門組織がこれを行うとした団体も3分の1程度(34.3%)に達した。

以上の結果から、大半の市・区は、行政評価の質を確保するために、評価実施者以外の主体が評価指標・目標値や評価結果の妥当性を判断する手続きを備えていることが確認された。妥当性の判断を行う主体は、庁内の推進部門と庁内の2次・3次評価者(主に1次評価者の上長)が中心であるが、評価指標や目標値については前者、評価結果については後者が担当する団体が相対的に多くなっている。また、都道府県では、第三者評価委員会等の専門組織を置いているところも3分の1程度存在するが、市・区においては一般的ではない。

言うまでもなく、行政評価の質を確保するための手続きを備えていることは望ましい。ただし、

表5 評価結果の妥当性判断の方法

判断する主体等	市・特別区		都道府県		全 体	
	回答	割合	回答	割合	回答	割合
庁内の推進部門が行う	81	32.4%	20	57.1%	101	35.4%
庁内の特別組織(タスク・フォース等)が行う	51	20.4%	4	11.4%	55	19.3%
庁内の2次・3次評価者が行う	123	49.2%	14	40%	137	48.1%
庁外のメンバーを主体とする専門組織(第三者委員会等)が行う	40	16%	12	34.3%	52	18.2%
明示的には行わない	62	24.8%	7	20%	69	24.2%
その他	4	1.6%	4	11.4%	8	2.8%
無回答	0	0%	0	0%	0	0%
合 計	250		35		285	

(注1) 行政評価制度を「導入済み」と回答した自治体のみについて集計。

(注2) 複数回答のため各項目(列)の回答割合の合計は100%にはならない。

(出所) 田中(2006a)

そのような手続きを備えていたとしても、それが必ず評価の質の向上に寄与するとは限らない。

実は、後述するように、「実態調査」では行政評価制度のさまざまな側面の状態に対する自治体の自己評価を尋ねており、その質問項目の中には、評価指標や目標値の質や評価結果の妥当性も含まれている。そこで、自治体が評価の質を確保するための手続きを備えているか否かが、評価指標・目標値の質や評価結果の妥当性とどのように関係しているかをクロス集計により調べてみた。

その結果、評価指標や目標値の妥当性を判断するための手続きを備えている自治体とそうでない自治体とでは、評価指標や目標値の質や妥当性に関する評価に大きな違いは認められなかった。一方、評価結果の妥当性を判断するための手続きを備えている自治体の方が、そうでない自治体よりも、評価結果の妥当性や信頼性については、より満足している傾向がみられた⁵。また、評価の質を確保するための手続きを備えていながら、評価指標・目標値の質や評価結果の妥当性に改善の必要性を認識している自治体も相当な割合に達している。

(4) 評価結果の予算策定への活用

自治体がどのような目的や手法で評価を実施す

るにせよ、行政評価制度と予算制度の間には必然的に関係性が生まれる。評価結果が何らかの予算措置を必要とするような対応(例えば、事務事業の拡大・縮小や職員の増員・減員など)を示唆する場合には、それを予算上に反映する必要があるからである。

原初的には、行政評価制度と予算制度とをそれぞれ個別に運用する中で、評価結果をケース・バイ・ケースで予算に反映していくことで対応が可能である。しかし、評価結果を系統的・包括的に予算に反映していこうとすれば、評価制度と予算制度とを有機的に結びつけて、両者を連動させていくが必要になる。事実、既に行政評価を予算策定過程に組み込んだり、評価制度と予算制度を一体的に運用したりしている自治体は少なくない。

この点について「実態調査」では、予算策定の各段階において、評価結果がどのような位置づけにあるかを質問した。その結果は表6に示すとおりである。予算策定を「各部門における予算要求作成の段階」「財政部門における予算査定の段階」「予算の査定後から最終決定までの段階」の3段階に分け、それぞれの段階において、評価結果の重要度を4つの選択肢(「不可欠」「ある程度重要」「あまり重要ではない」「ほとんど利用されず」)の中から選んでもらった結果である。

この結果から、市・区の大多数が、部門の予算要求作成の段階や財政部門の予算査定の段階において、評価結果を何らかの形で利用していることが明らかになった（予算要求作成段階では約83%、予算査定段階では約74%の自治体が、評価結果を「不可欠」または「ある程度重要」と回答）。ただし、これらの段階において評価結果が「不可欠」であるとした団体はそれほど多くはなく、半数程度の都市自治体が評価結果を「ある程度重要」と回答したに留まっている。都道府県においても回答の傾向は同様であった。

この調査結果だけからでは、予算策定プロセスに行政評価がどの程度系統的に連動しているかは依然として明らかではないものの、予算要求の作成段階や予算査定の段階において、評価結果が「不可欠」ではないものの、「ある程度重要」な判断材料の1つとして利用されているという実態が確認された。

4. 都市自治体における行政評価の成果と問題点

(1) 行政評価の実施による成果

自治体が行政評価制度を導入・実施した成果に

ついては、これまでほとんど解明されていない⁶。しかし、自治体における評価の導入が進んだことから、行政評価を実施した成果がそろそろ問われるべきである。そこで「実態調査」では、行政評価制度を導入している自治体において、どのような成果が（評価の推進担当者によって）認識されているかを質問した。

この質問においては、想定される回答をあらかじめ選択肢（10個）として設定し、それぞれの選択肢について、（そのような成果が）「大いに当てはまる」「当てはまる」「どちらとも言えない」「あまり当てはまらない」「ほとんど当てはまらない」の5つの中から1つを選んでもらった⁷。各選択肢が示す成果を認識している（「大いに当てはまる」または「ある程度当てはまる」）と回答した自治体数の割合を示したのが図2である。

市・区についてみると、成果を認識した自治体が特に多かったのは、「有効性の低い事務事業の廃止・削減」（62%が成果を認識）と「企画立案や予算策定の際に『成果』の観点で検討」（57.2%が成果を認識）の2項目であった。これらに次いで多かったのは、「個別の事務事業の効率性向上」（47.2%が成果を認識）「議会における評価結果への関心向上」（46.8%が成果を認識）「個別の事務事業の有効性向上」（44.8%が成果を認識）「事務事業間の重複解消」（43.6%が成果を認

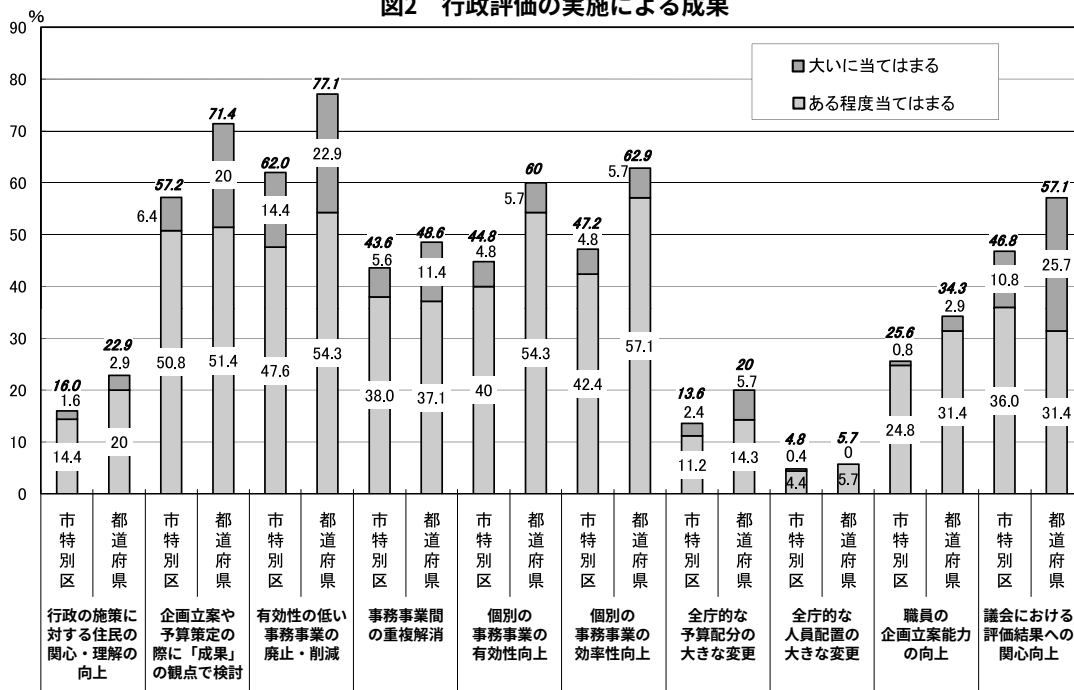
表6 評価結果の予算策定プロセスにおける活用

		不可欠＋ある程度重要						あまり重要ではない		ほとんど利用されず		無回答		合 計	
				不可欠		ある程度重要									
		回答	割合	回答	割合	回答	割合	回答	割合	回答	割合	回答	割合	回答	割合
市・特別区	各部門における 予算要求作成の段階	29	82.9%	10	28.6%	19	54.3%	5	14.3%	1	2.9%	0	0%	35	100%
	財政部門における 予算査定の段階	26	74.3%	5	14.3%	21	60%	5	14.3%	3	8.6%	1	2.9%	35	100%
	予算の査定後から 最終決定までの段	13	37.1%	4	11.4%	9	25.7%	11	31.4%	9	25.7%	2	5.7%	35	100%
都道府県	各部門における 予算要求作成の段階	175	70%	61	24.4%	114	45.6%	41	16.4%	34	13.6%	0	0%	250	100%
	財政部門における 予算査定の段階	158	63.2%	31	12.4%	127	50.8%	49	19.6%	41	16.4%	2	0.8%	250	100%
	予算の査定後から 最終決定までの段階	116	46.4%	22	8.8%	94	37.6%	62	24.8%	71	28.4%	1	0.4%	250	100%
全 体	各部門における 予算要求作成の段階	204	71.6%	71	24.9%	133	46.7%	46	16.1%	35	12.3%	0	0%	285	100%
	財政部門における 予算査定の段階	184	64.6%	36	12.6%	148	51.9%	54	18.9%	44	15.4%	3	1.1%	285	100%
	予算の査定後から 最終決定までの段階	129	45.3%	26	9.1%	103	36.1%	73	25.6%	80	28.1%	3	1.1%	285	100%

（注）行政評価制度を「導入済み」と回答した自治体のみにについて集計。

（出所）田中（2006a）

図2 行政評価の実施による成果



(注1) 行政評価制度を「導入済み」と回答した自治体のみについて集計。

(注2) 太字斜体の数字は、「大いに当てはまる」と「ある程度当てはまる」を合計した数値。

(出所) 田中 (2006a)

識」といった項目であった。事務事業に関連する成果が認識されているのは、先にみた行政評価の導入目的において、事務事業レベルにおける改善に関連する目的を挙げた自治体が多かった事実と整合している。

一方、成果があまり認識されていないのは、「全庁的な人員配置の大きな変更」(4.8%が成果を認識)「全庁的な予算配分の大きな変更」(13.6%が成果を認識)「行政の施策に対する住民の関心・理解の向上」(16%が成果を認識)「職員の企画立案能力の向上」(25.6%が成果を認識)といった項目であった。

このうち「行政の施策に対する住民の関心・理解の向上」という成果は、行政評価の導入目的として大半の自治体を選んだ「住民等の行政以外の主体に対して行政の活動状況を明らかにする」という目的に対応している。この結果から、自治体は行政評価によってその活動状況を住民に対して示しているものの、現時点では行政から住民への

一方的な情報提供に留まっている状況がうかがえる。

「全庁的な予算配分の大きな変更」と「全庁的な人員配置の大きな変更」という成果は、やはり導入目的で挙げた「全庁レベルの予算配分の変更」と「全庁レベルの人員配置の変更」にそれぞれ対応している。行政評価の導入目的としてこれらを選択した自治体が少なかったことから、これらの成果が認識されていないのは、ある意味では当然の結果である。

最後に「職員の企画立案能力の向上」という成果は、行政評価の導入に伴って自治体が期待する効果のひとつとして言及されることが多い。しかし、2～3割の自治体がこの成果を認識するに留まっており、自治体側の期待を裏切る結果となっている。

この結果の解釈については引き続き検討が必要であるが、行政職員が事務事業を評価する活動に必要とされる能力と事務事業の企画立案作業に必

要な能力との間には、関連性があるとはいえ、少なからぬ距離があることを示唆している可能性がある。

以上では、都市自治体の回答結果をみたが、都道府県の回答の傾向もほぼ同様である。ただし、全般的に都道府県の方が成果を認識している団体の割合が高くなっている。

また、図2には示されていないが、選択肢として挙げたいかなる項目についても成果を認識していない自治体は32団体（28市・区、4都道府県）であり、全体の11%であった。

なお先に指摘したように、これらはあくまでも行政評価の推進担当者の認識を把握した結果である。したがって、自治体における行政評価の成果を正確に反映しているとまで断言することができない。とはいえ、図2が示すように、成果の項目によって自治体の回答率に大きな差が出たということは、（たとえ推進担当者の回答にバイアスがあるとしても）成果の一定の傾向を反映したものと理解することができよう。

(2) 成果と実施期間の関係

以上から、行政評価制度を導入している市・区や都道府県においては、いくつかの側面において成果が認識されていることが示唆された。さらに、市・区について、行政評価の成果と実施期間との関係をみたのが図3である。この図には、図2と同様、行政評価の成果を代表する10の選択肢について、それぞれの成果を認識している（「大いに当てはまる」あるいは「ある程度当てはまる」と回答した）自治体の割合が行政評価の実施期間（行政評価制度を導入後の期間）別に示されている。

この図から明らかなように、図2で成果を認識した自治体が多かった項目（「企画立案や予算策定の際に『成果』の観点で検討」「有効性の低い事務事業の廃止・削減」「個別の事務事業の有効性向上」「個別の事務事業の効率性向上」「議会における評価結果への関心向上」）においては、自治体の実施期間が長くなるほど成果を認識する割合が増加する傾向がある⁸。

一方、成果の認識度合いが低い項目（「行政の施策に対する住民の関心・理解の向上」「全庁的

な予算配分の大きな変更」「全庁的な人員配置の大きな変更」「職員の企画立案能力の向上」）においては、成果と実施期間の間に明確な関係は認められない。これらの項目においては、評価の経験を積んでも、大きな成果の向上が期待できないということが示唆されている。

(3) 行政評価制度の状態についての認識

行政評価制度が期待されるような機能を果たすためには、それが良好な状態で維持・運用される必要がある。そこで「実態調査」では、評価の推進担当者が自団体の行政評価制度の状態についてどのように認識しているかを把握しようと試みた。ここで「行政評価制度の状態」とは、制度の仕組み（制度設計）の良し悪しだけでなく、制度を利用して行われる評価の質、制度の運用体制、さらには制度の運用によって得られる評価結果の利用状況を含む幅広い意味で用いている。

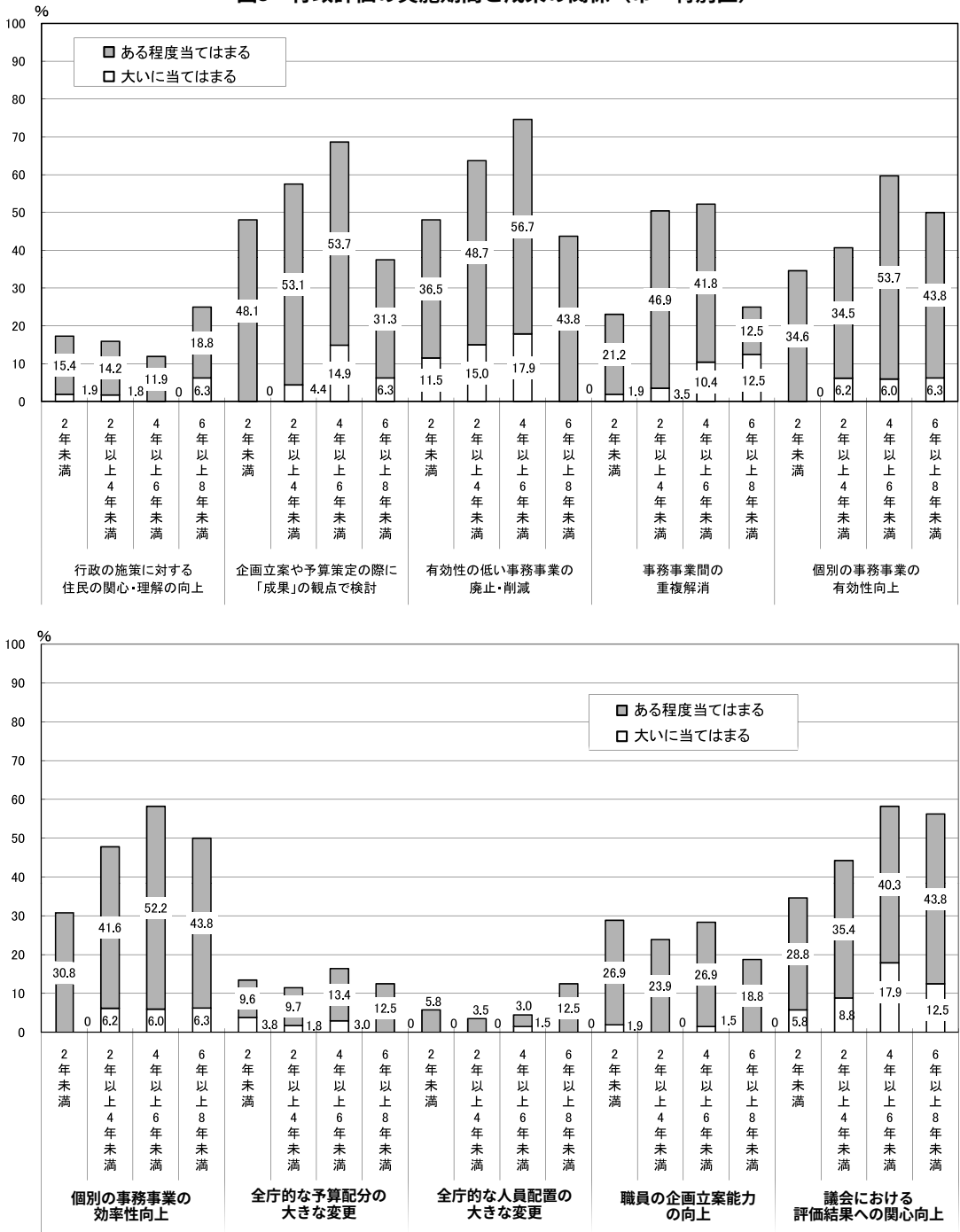
これらを具体的に把握するために、「実態調査」においては、「制度」「評価」「運用体制」「結果の利用」の4つの側面別に具体的な項目を選択肢として示し（表7）、それぞれの事項について、各自治体（における評価の推進担当者）がどのように認識しているかを4つの選択肢（「満足できる水準にある」「ある程度満足できる水準にある」「どちらとも言えない」「多少の改善が必要である」「大幅な改善が必要である」）の中から選んでもらった。また、行政評価制度の全体的な状態（総合評価）についても各自治体の認識を質問した。

図4には、この質問に対する回答のうち「満足できる」あるいは「ある程度満足できる」と回答した自治体の割合を示した。まず行政評価制度の全体的な評価（総合評価）からみると、市・区の2割強（23.2%）が「満足できる」あるいは「ある程度満足できる」と回答するに留まった。一方、都道府県においては40%が「ある程度満足できる」と回答しているので、行政評価の全体的な状態についての認識は、都道府県の方が良好である。

側面別にみると、「制度」面（「評価制度の仕組み」と「評価対象の妥当性」の2項目）は他の側面に比べて満足している市・区が多い。

「評価」面（5項目）は、項目によって評価が分

図3 行政評価の実施期間と成果の関係（市・特別区）



(注) 行政評価制度を「導入済み」と回答した市・特別区のみについて集計。

(出所) 田中（2006a）

表7 「実態調査」における行政評価制度の状態の選択肢

行政評価制度の主要側面	主要側面別の具体的項目(選択肢)
a. 制 度	a-1. 評価制度の仕組み
	a-2. 評価対象の妥当性
b. 評 価	b-1. 評価指標の質
	b-2. 目標値の妥当性
	b-3. 使用するデータの妥当性・信頼性
	b-4. 評価結果の妥当性・信頼性
	b-5. 職員の評価に関する理解や能力
c. 運用体制	c-1. 推進部門の能力(体制とノウハウ)
	c-2. 庁内の協力体制
	c-3. 評価のチェック体制
d. 結果の利用	d-1. 職員による評価結果の活用
	d-2. 各部門における評価結果の利用
	d-3. 全庁的な意思決定(重点政策の決定等)における評価結果の利用
e. 総合評価(行政評価制度の全体的な評価)	

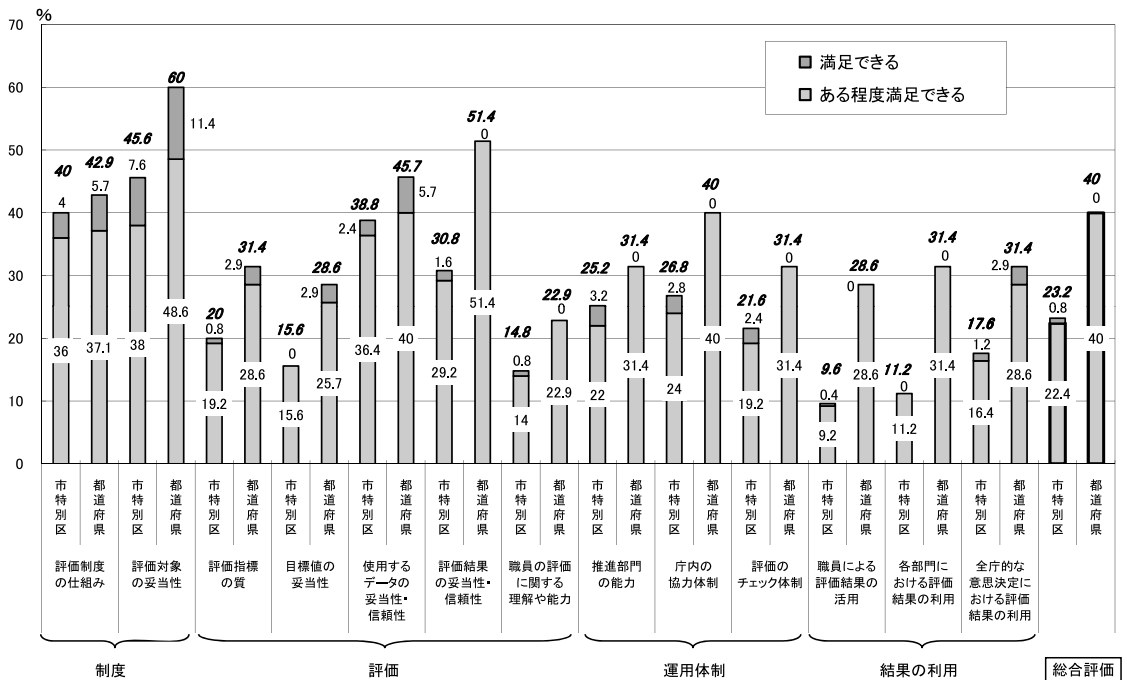
(出所) 筆者作成

かれており、市・区で満足を示した自治体が相対的に多かったのは「使用するデータの妥当性・信頼性」と「評価結果の妥当性・信頼性」であった。一方、「評価指標の質」「目標値の妥当性」「職員の評価に関する理解や能力」といった項目について満足を示した自治体は2割以下であった。

「運用体制」面（「推進部門の能力」「庁内の協力体制」「評価のチェック体制」の3項目）については、項目間の回答の差異はあまり大きくなく、市・区の2～3割がこの側面の項目に対して満足りていると回答した。

最後に「結果の利用」面については、都道府県と市・区の回答の傾向に大きな差が見られた。都道府県においては、この側面のいずれの項目とも30%前後の団体が満足りていると回答したが、市・区で満足りていると回答した自治体は、回答率が最も多かった「全庁的な意思決定における評価結果の利用」という項目においても2割を下回

図4 行政評価制度の各側面の状態に対する満足度



(注1) 行政評価制度を「導入済み」と回答した自治体のみについて集計。
(注2) 太字斜体の数字は、「満足できる」と「ある程度満足できる」を合計した数値。
(出所) 田中 (2006a)

った。

以上を総合すると、行政評価制度の状態（についての認識）は都道府県の方が市・区よりも良好であり、その差は「結果の利用」面の認識に特に顕著に現れている。行政評価制度の総合評価においても、都道府県の4割が満足感を示しているのに対し、市・区においては4分の1程度が満足しているに留まっている。

先に図2でみた行政評価の実施による成果では、都道府県と市・区の認識に大きな傾向の違いがなかったことから、行政評価制度の状態に関する両者の認識にこれほど顕著な差が観察されたことは注目に値する。

この点については1つの解釈の可能性を指摘しておきたい。それは行政評価制度の状態と行政評価制度の成果が必ずしも直接的に関係していないという可能性である。

市・区の回答結果が示すように、評価指標や目標値の質が低く、評価結果がいかなるレベルにおいても十分に活用されていないとすれば、具体的な成果につながるとは考えにくい。しかし、市・区の6割以上が「有効性の低い事務事業の廃止・削減」を成果として認識しているという事実は、現在の行政評価制度が、たとえ評価の質や評価結果の利用が不十分であっても、一定の成果が認知されやすい制度設計になっていることを示唆している。換言すれば、評価指標を利用した業績測定の結果が何であれ、それとは切り離して事務事業単位の改廃や改善が実施されている可能性がある。

このような推察が正しいとすれば、純粋な評価論の立場からは甚だ残念な実態である。一方、評価結果を直接利用しないとしても、行政評価制度の枠組みを利用して事務事業レベルを中心とした改革や改善が進んでいるとすれば、（それらが適切である限りは）特に問題視しないという立場もあろう。

（4）行政評価制度の状態と実施期間の関係

図4でみた行政評価制度の状態についても実施期間との関係を確認した。その結果は図5に示されている。

図5は、図4に示したのと同じ質問に対する市・区の回答結果であるが、図4と異なるのは、行政評価制度の各項目について「改善の必要を認識している」（つまり、各項目に対して「大幅な改善が必要」または「多少の改善が必要」と回答した）自治体の割合を示している点である。

これによれば、先にみた行政評価制度の状態に関する満足度が低かった項目（「評価指標の質」「目標値の妥当性」「職員の評価に関する理解や能力」「職員による評価結果の活用」「各部門における評価結果の利用」など）においては、改善の必要性を認識している自治体の割合が総じて高くなっている。しかも、自治体の実施期間が長くなっても、改善の必要性を認識する団体の割合はほとんど減少していない。つまり「評価指標の質」「目標値の妥当性」「職員の評価に関する理解や能力」といった評価活動に直接関わる項目については、行政評価制度の運用経験がその改善にあまり貢献していないことが示唆されている。

一方、「職員による評価結果の活用」や「各部門における評価結果の利用」という「結果の利用」に関する項目についても、実施期間との明確な関係は観察されなかったが、これは恐らく行政評価制度の制度設計に起因するものであろう。つまり、評価制度が構造的な問題を抱えているために、運用期間が長くなっても評価結果の利用が進んでいない可能性が高い。

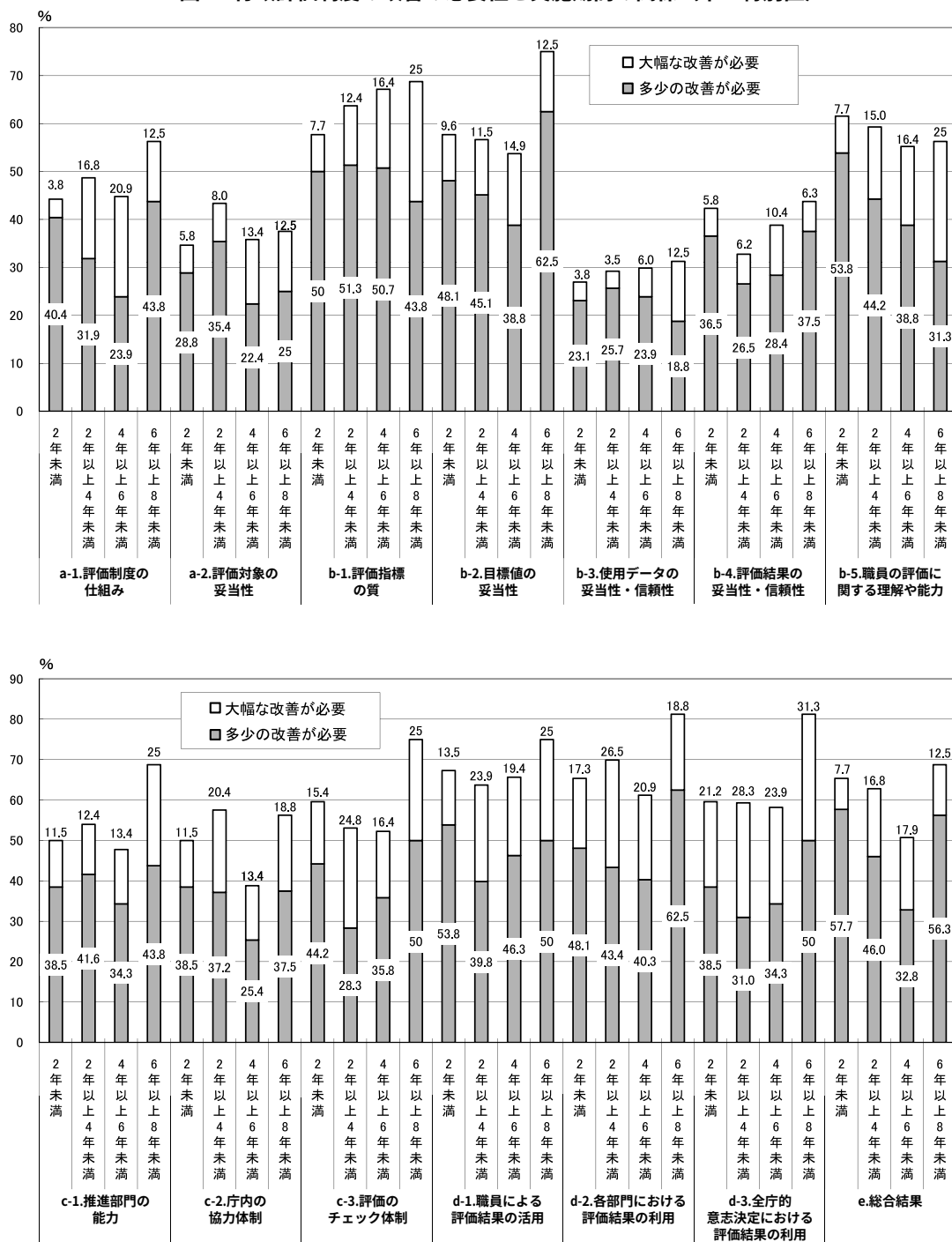
（5）行政評価制度の問題点

以上の分析をもとにすると、都市自治体の行政評価制度については、次のような問題点を指摘することができる。

まず行政評価制度の全般的な状態について改善の必要性を感じている都市自治体（の評価推進担当者）が多く、都道府県と比較しても、市・区の行政評価制度の状態は良好であるとは言えない。

行政評価制度の側面別にみると、全般的に改善の必要性が認識されているが、評価活動の質に関わる側面（評価指標、目標値、職員の理解や能力）や評価結果の利用面（職員や各部門による利用）においては、改善の必要性を認識している自治体が特に多くなっている。

図5 行政評価制度の改善の必要性和実施期間の関係（市・特別区）



（注）行政評価制度を「導入済み」と回答した市・特別区のみについて集計。

（出所）田中（2006a）

さらに注目すべきは、経験を積んだ自治体においても、行政評価制度の状態が大きく改善していないことである。この事実は、ただ単に行政評価の経験を積むだけでは、行政評価制度の問題点を解決することが困難である可能性を示唆している。

一方、行政評価制度の状態の良し悪しが、成果の有無に直接影響していない可能性も示唆された。つまり、評価活動の質や評価結果の内容とは直接の関係なく、事務事業の改廃等がなされている可能性である。このことは、現実主義的な立場からは容認されるかもしれないが、行政評価のありべき姿ではないことは言うまでもない。

5. 結論

(1) 都市自治体の評価の実態

本稿では、筆者が実施した「実態調査」の結果をもとにして、都市自治体の行政評価の実態を明らかにしようと試みた。その結果、これまで経験的（あるいは断片的）に認識されていたものの、確たる実態がつかめていなかった点について傍証が得られたほか、いくつかの点では今回新たに示唆が得られた。

まず、これまで経験的・断片的に認識されていたものの、今回の分析によって追認することができたのは、以下の諸点である。

- ①都市自治体における行政評価の実施期間は総じて都道府県よりも短く、大半の自治体は4年未満の実施経験に留まっている。
- ②行政評価の導入目的として事務事業レベルの改善とアカウントビリティ関係の目的を設定している自治体が圧倒的に多い。
- ③評価結果は、予算策定プロセス（特に予算要求の作成時と予算査定時）において重要な要素として利用されているが、不可欠な要素として認識されるには至っていない。
- ④大半の自治体が行政評価の質を確保するための手続き（推進部門によるチェックや2次・3次評価など）を備えている一方、こうした手続きを

持たない自治体も少なからず存在する。

- ⑤行政評価の質の良し悪しに関わらず、評価の結果として事務事業の改廃等が行われている可能性がある。

一方、今回の分析によって示唆された点のうち、本稿の観点に照らして特に注目すべきは、以下の点である。

- ⑥都市自治体の行政評価制度は一定の成果を上げているが、成果が上がっている分野は職員の意識変化や事務事業レベルの改善に限定されている。
- ⑦一方、全庁的な資源（予算・人員）配分面では、成果を認識している自治体は少ない。
- ⑧これらの側面においては、実施期間の長い自治体においても成果がほとんど認識されておらず、現行の事務事業評価を中心とする行政評価制度が全庁的な資源配分面では有効でないことを示唆している。
- ⑨評価指標や目標値の質や職員の評価能力について問題を抱えている都市自治体が少なくなく、しかも行政評価の経験が蓄積されても、これらの問題が解決される可能性が低い。

(2) 自治体評価の今後の方向性

これまでの検討を踏まえて、都市自治体を中心とする自治体評価が今後めざすべき方向について筆者の私見を示したい。

第1に、多くの自治体が採用してきた事務事業評価型のアプローチに限界が見えてきたのではないかと考えられる。

確かに事務事業レベルの改善等の成果は認識されているものの、実態を仔細に分析すると、質の良い評価の結果としてそのような成果が上がっているのではないことが示唆されている。言わば「結果ありき」で評価が実施されている可能性があり、そのような行政評価制度は、早晚自治体内部で職員が真剣に取り組む対象ではなくなるであろう。最近「事業仕分け」に取り組む自治体が登場している背景には、自治体が従来の事務事業評価型のアプローチに限界を感じていることも関係

しているのではないかと考えられる。

事務事業評価型の評価制度がすぐに不要になるわけではないとしても、これに替わる評価のアプローチの必要性が高まっていくことが予想される。

第2に、上記の方向性に関連して、全庁的な行政資源、特に予算の配分に有効な行政評価制度への志向性がますます強まっていく可能性がある。

既に指摘したように、評価結果を具体的に実現していくためには、何らかの形で予算上に反映していくことが必要であり、この方向性を推し進めていけば、行政評価制度と予算制度とを有機的に連動させることに行き着く。

しかし、本稿の分析が示唆するように、現行の行政評価制度は、全庁的な資源配分の変更にはほとんど有効ではないことから、予算制度と連動させることを前提とした新たな行政評価制度の開発が必要になる。事実、諸外国や日本の中央府省において近年焦点となっているのは、まさに予算制度と評価をいかに連動させるかという点である¹⁰。

これについては、今のところ画期的なアプローチが開発されたわけではないため、当面は試行錯誤を続けながら、より良い方策を模索していくことになるう。

第3に、やはり多くの自治体に取り組んできた業績測定型のアプローチの難しさも明らかになっている。そもそも業績測定型のアプローチは、プログラム評価に代表されるような、より厳密性の高い評価手法よりも簡便に実施できることから、自治体の間に広く普及してきたという事情がある。

しかし、業績測定に不可欠な評価指標や目標値の質を向上させることが決して容易ではないことが「実態調査」によって明らかになっている。また評価活動を実施する職員の理解や能力の向上が極めて難しいことも分析結果が示唆している。

このことから、評価に取り組む自治体は、業績測定が簡便な手法であるとの先入観を払拭した上で、これに対して真剣に取り組んでいくか、あるいは、別の手法を志向していくかの選択を迫られることになるう。

以上を総合すると、自治体評価の今後の方向性としては、「脱事務事業評価」「予算と評価の連動」

「脱業績測定」というキーワードに集約することができる。

ただし、これまでの自治体評価のあり方が直ちにこれらの方向性にとって替わられるわけではなく、これらのキーワードに象徴される新しい志向性が徐々に浸透していく可能性があると理解すべきである。

以上はあくまで可能性として指摘したものであり、これらの方向性を実現するかどうかは、自治体自身の認識と取り組みのあり方のほか、自治体の要請に応えることのできるような新しい行政評価の手法やアプローチが提示されるかどうかにもかかっている。その意味では、評価研究に関わる研究者の貢献がますます期待されることになるう。

なお、今回は探索的研究として自治体における行政評価の実態を概観することに留まったが、自治体をグループ分けして行政評価の実態の差異を検出することや、行政評価を導入していない自治体を研究対象とすること等については、今後の研究課題としたい。

謝辞

本稿の完成にあたり、査読者の方から丁寧かつ有益なご指摘を多数いただきました。この場を借りて深謝いたします。

注記

- 1 本稿では、地方自治体が実施する行政評価を便宜的に「自治体評価」と表現する。
- 2 総務省の調査以外に代表的なものとしては、三菱総合研究所が毎年実施している「地方自治体における行政評価への取り組みに関する実態調査」がある。
- 3 特に本稿が重視する行政評価の成果や問題点については、回答者の主観が反映されやすい。
- 4 これは都道府県と市・特別区を併せた集計結果である。
- 5 評価指標や目標値については、これらの質を確保するための手続きの有無が質に関する認識の違いにあまりつながっていないのに対して、評価結果については、手続きの有無によって評価結果の妥当性に対

- する自治体の認識に大きな差が出ているという結果は、評価指標や目標値が形骸化している可能性を示唆している。この点については4 (3) で再び述べる。
- 6 行政評価の実施による効果を行政サービス度という総合指数の変化によって把握しようと試みた研究として横山 (2006) があるが、本稿のように、行政活動の側面別の成果を把握しようと試みた先行研究は存在しない。
 - 7 選択肢を設定する際には、前述した行政評価の導入目的の選択肢との対応関係に配慮した。
 - 8 ただし、実施期間が「6年以上8年未満」の自治体は、成果を認識する自治体の割合があまり高くない。この点についてはいろいろな解釈がありうるが、この区分に該当する自治体の数が多くない (16団体) 点には留意が必要である。
 - 9 この点に加えて実施経験が長くなるほど、行政評価制度に対する自治体自身の期待値が高くなり、その状態についてより厳しく評価するという効果が働いている可能性もある。
 - 10 例えば米国では、GPRA (政府業績結果法) の運用の一環として、評価体系と予算体系の関連性を高めてきたほか、ブッシュ政権の下では、PART (Program Assessment Rating Tool) というツールを導入して、プログラム別の評価結果と予算配分とを連動させている。一方、日本の中央府省では、2005年12月に政策評価と予算・決算の連携を強化する方針が示され、各府省において両者の体系を連結する取り組みが進められている。

参考文献

- 小野達也 (2002) 「地方自治体の行政評価システムの課題と展望」、『日本評価研究』、2 (1) : 29-38
- 窪田好男 (2005) 『日本型政策評価としての事務事業評価』、日本評論社
- 島田晴雄・三菱総研 (1999) 『行政評価—スマート・ローカル・ガバメント—』、東洋経済新報社
- 総務省 (2005) 『新地方行革指針による地方行革の推進』
- 総務省 (2006) 『地方公共団体における行政評価の取組状況』
- 田中啓 (2004) 『自治体の行政評価—組み込まれた不機能とそれへの対応—』、富士通総研経済研究所 Research Paper Public Series No. 2
- 田中啓 (2005a) 「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)
- 田中啓 (2005b) 『自治体評価にみる本質的課題とその解決方法に関する考察』、日本評価学会・第6回全国大会・報告論文
- 田中啓 (2006a) 『行政評価の実態に関する調査・調査結果報告書』
- 田中啓 (2006b) 『地方自治体の行政評価：10年間の実践を診断する』、日本評価学会・第7回全国大会・報告論文
- 田中啓 (2007) 「行政評価の実態に関する調査—調査結果の要約—」、『評価クォーターリー』、1 : 38-46
- 古川俊一 (2001) 「地方自治体の行政評価」、『日本評価研究』、1 (1) : 25-37
- 古川俊一・北大路信郷 (2004) 『新版公共部門評価の理論と実際』、日本加除出版株式会社
- 三菱総合研究所 (2006) 『地方自治体における行政評価への取り組みに関する実態調査』
- 山谷清志 (2006) 『政策評価の実践とその課題：アカウンタビリティのジレンマ』、萌書房
- 横山麻季子 (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
- De Lancer Julnes, P., & Holzer, M. (2001) . Promoting the utilization of performance measures in public organizations: an empirical study of factors affecting adoption and implementation. *Public Administration Review*, 61 (6) , 693-708.
- Governmental Accounting Standards Board (GASB) (2002) . *Performance Measurement at the State and Local Levels: A Summary of Survey Results*.
- Hatry, H. (1999) . *Performance measurement: getting results*. The Urban Institute Press, Washington, D.C.
- Hatry, H. (2006) . *Performance measurement: getting results. 2nd Edition*. The Urban Institute Press, Washington, D.C.
- King, G., Keohane, R.O., & Verba, S. (1994) . *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton University Press, Princeton, NJ.
- Osborne, T., & Gaebler, T. (1993) . *Reinventing Government: How The Entrepreneurial Spirit Is Transforming The Public Sector*. Addison-Wesley, Reading, MA.

Tanaka, H., & Ono, T. (2002) . Performance Measurement Utilization in Japan: A Critical View. A Paper presented at the Annual National Conference of the ASPA.

Tanaka, H. (2006) . Current State of Performance measurement in Japan: Are They Properly Functioning? A paper presented at the EGPA conference 2006 in

Milan.

Tanaka, H. (2007) . Factors Affecting The Performance of Performance Measurement Systems: An Analysis of The Japanese Practice. A paper presented at the EGPA conference 2007 in Madrid.

(2008.2.1受理)

参考資料 「行政評価の実態に関する調査」の概要

付表1 実施方法

調査対象：2006（平成18）年2月末時点で国内に存在する全ての都道府県・市・特別区

都道府県	47団体	} 合計 833団体
市・特別区	786団体	

調査方法：郵送調査法（A4版4ページの調査票を各団体に郵送）

調査期間：2006（平成18）年3月3日～28日

想定される回答者：各団体において、行政評価制度の導入・運用面で主導的な役割を果たしている職員（推進部門の長や担当者）

質問項目：行政評価の導入状況、取り組みの契機、導入時の体制、参考にした自治体、導入目的、評価対象、記載を求める評価指標、評価指標・目標値・評価結果の妥当性判断の方法、予算策定プロセスにおける評価結果の活用状況、導入の成果、行政評価制度の状態に関する認識、課題や問題点

※調査票は無記名とし、団体名や回答者の所属・氏名の記入を求めなかった。

付表2 回答状況

自治体区分	発送数	回答数	回収率（％）
都道府県	47	36	76.6
市・特別区	786	528	67.2
政令指定都市	14	9	64.3
特別区	23	16	69.6
上記以外の市	749	503	67.2
合 計	833	564	67.7

（注）寄せられた回答は全てが有効な回答であった。

Performance Measurement in City Governments in Japan: The Current State Ten Years after Its Introduction

Hiraki Tanaka

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

Abstract

Ten years or so have passed since Japanese local governments began to incorporate performance measurement and related activities into their management practice. This study is an exploratory study of which objective is to grasp the current state of performance measurement in the local government in Japan. In doing so, this study tried to examine the state of performance measurement in the city governments (including special wards in Tokyo) by utilizing results of a survey conducted by the author. In addition, special attention was given to the outcomes of performance measurement implementation and the problems of performance measurement systems. As a result of this study, significant findings were obtained such as the following; current performance measurement systems are effective in changing the mindsets of public employees and in pursuing improvements at project level, while they are not very useful for changing the government-wide resources allocation; and a significant number of city governments are not free from problems in terms of the quality of evaluation activities and the evaluation capacity of public employees. Based on the analysis of this study, it is suggested that the key trends of performance measurement at local level in Japan might be “departure from project level evaluation,” “linking performance measurement with budget process,” and “departure from performance measurement.” Whether such trends will be realized or not, depends on how local governments find and react to the problems of their own performance measurement systems and what kind of work researchers in this field would be able to contribute.

Keywords

local government evaluation, performance measurement,
results of performance measurement, linkage between performance measurement and budget

【研究論文：依頼原稿】

公共政策学・政策評価論・日本型政策評価

窪田 好男

神戸学院大学 法学部

kubota@law.kobegakuin.ac.jp

要 約

公共政策学は今まさに成長し、定着しつつある新たな学問分野であり、社会的に重要な役割を期待される学問分野である。その意図するところは、ポリシーマインド（政策学的思考）を確立し、科学・学問によってよりよい社会を実現することである。

その公共政策学の中で、公共政策の評価を研究対象とする政策評価論は、公共政策学原論、すなわち公共政策学の中核の一部を担うと考えられている。ところがわが国における政策評価の理論と実践は期待に届いてきたとは言い難い。そうなった原因は、わが国の政策評価の実践の多くが三重県の事務事業評価システムをベースとしその発展上にあるシステムを誤用していること、およびわが国の政策評価研究の多くが実際に導入された制度とその運用の批評に終始していることに求めることができる。

事務事業評価システムをベースとする政策評価システムの役割は、事務事業のカットによる財政状況の改善などではなく、行政職員の意識改革にとどまるものでもなく、公共政策のデザインのための情報提供と再定義されなくてはならない。

キーワード

公共政策学、事務事業評価システム、政策評価論の課題、政策評価の目的、政策デザイン

1. はじめに

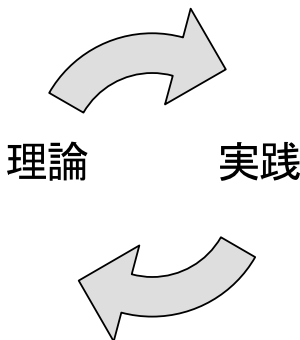
三重県に事務事業評価システムが導入されたのは1997年である。それから10年以上が経過し、都道府県や政令指定市、中核市や特例市など大規模な自治体の多くが類似のシステムを導入している。その導入や運用の実態については、本学会でも多くの業績が蓄積されつつある。本稿では、日本型政策評価¹である事務事業評価について、あらためて政策評価論におけるその位置づけを明らかに

かにしたい。さらには、政策評価論が政策デザインも包含するか、少なくとも関わるということで、興隆しつつあるディシプリンである公共政策学の中核を担うことを示したい。

わが国の評価については、短期的で可視性の高い改革効果の観点から、あるいは政策評価の実施を担う行政職員の負担といった観点から論じられることが多かった。学問の体系の中で政策評価がどのような位置を占めるかに行った議論は多くはなかった。政策評価や公共政策といった学問分野

では、理論と実践が適切な相互作用関係を構築することが、それぞれの発展に必要である。政策評価について、負担は多大であるにもかかわらず目に見えるような短期的効果が少ないという見解が広まりつつある現在、学問の体系における政策評価の位置づけを再確認し、理論と実践の相互作用を活性化することが強く求められているのではないだろうか²。

図1 理論と実践の相互作用



(出所) 筆者作成

以下、2では、公共政策学というディシプリンについて考察し、政策評価論との関係を明らかにしたい。3では、政策評価論の展開を概観し、公共政策学と政策評価の関係を観察する観点、および日本型政策評価としての事務事業評価を観察する観点を得たい。4では公共政策学の政策評価論において日本型政策評価がどのようなポジションを占めるかを明らかにしたい。

ところで、政策評価論は公共政策学というディシプリンの中核的位置を占めるというのが本稿の主張であるが、政策評価を含めた評価自体が一つのディシプリンなのではないかという考え方もある。研究書も増えているし、大学の学部や大学院の講義科目ともなっている（山谷 2007、p. 2）。日本評価学会という研究者や実務家が集まる場としての学会も存在する。アメリカ合衆国ではディシプリンとして確立しているという指摘もある（龍・佐々木 2004、p. 17）。しかし、評価を、法学や政治学、経済学、経営学、社会学などと並ぶディシプリンと見なすことにはためらいを感じざるを得ない。3で詳述するが、評価は応用社会

科学（applied social science）と自己規定してきた。応用社会科学といっても、実際の所は応用社会調査と限定的に領域を定めた方が実態に近く、その意味するところは、政策形成過程の各段階で社会調査を行いその結果を政策改善に活用するということである。要点は、既存のディシプリンの手法やそれらによって蓄積されてきた学知を政策過程に応用・活用するということであり、そうであるとするならば評価が独自のディシプリンとは言い難いと思われるのである。

ディシプリンとなるためには、研究者や実務家が集う場としての学会があるだけでは不十分であり、共通の学知を開発・蓄積・体系化しようというパトスが学会構成員に共有されること、研究者や実務家だけでなくそれを学ぶ学生まで含めてアイデンティティを持つこと、ディシプリンを通して習得できる知識や技能の中身についての共通の了解ないしイメージが共有されていることの3点が必要であるとされる（足立 2005、pp. 12-14）。これらの点からすると、政策評価は未だ独自のディシプリンとして成立しているとは言い難いのではないだろうか。

政策評価が公共政策学のディシプリンの中核に位置づけられるためには、理論と実践に一定の再構築が求められる。3で詳述するが、政策評価は、議題設定から終了に至る政策形成の全段階に関与し、公共政策や政策過程を改善しようとする評価へと進化している。その展開の中で、従来からの応用社会科学という自己規定を堅持することは不可能であるし適切でもない。価値の問題や公共哲学の問題など政策をデザインする際に欠くべからざる要素を取り入れて自らを再規定することが必要とされる。

2. ディシプリンとしての公共政策学と政策評価論の3つの課題

公共政策学については、他にも政策科学、総合政策学、政策学といった呼び方もある。政策科学が主にアメリカ合衆国で発展してきた“Policy Sciences”の直訳であるとか、総合政策学は政策に関する縦割りの改革を主眼とするとか、政策学

は公共政策やその部分集合である政府政策だけではなく企業の政策³も扱うといったそれぞれの特徴はあるが、完全に別のディシプリンということはない。ポリースクール、あるいは政策系と総称されることもあるが（政策分析ネットワーク編 2003）、全国的な学会の名称にもなり定着しているのは公共政策学である。

名称の相違以上に問題となるのが、そもそも公共政策学は独自のディシプリンといえるかという問題である。ディシプリンでないなら、その中に政策評価の理論と実践を位置づけることは不可能だからである。

政策科学や総合政策学は、政策科学が原語である“Policy Sciences”を忠実に和訳すれば政策諸科学となることに端的に表れているが、自らを応用社会科学と規定する。独自のディシプリンを構築するというよりは、経済学や経営学、政治学、法律学、社会学といった既存のディシプリンの手法や学知を応用し、よりよい社会を実現する政策提言をまとめることを重視する。さらに、総合政策学については、エリートの役割を重視し、政策分析や評価なども決定作成権者（decision maker）への情報（intelligence）^{インテリジェンス}提供手段と見なす傾向がある（政策分析ネットワーク編 2003、第I部2）。もちろん、既存のエリートである政治家や官僚、マスコミ等をいたずらに擁護するのではなく、改革を志向する運動のロジックとしては、それらを新たな政策系の人材によって代替することを意図しているわけであるが、いずれにせよ、エリートによる社会科学の手法や学知の応用による社会改革—しかも比較的短期かつ即効的な改革—を意図していることには違いない。

他方、公共政策学と称する場合には、多かれ少なかれ、独自のディシプリンであるという自己規定をしている場合が多いように思われる。独自のディシプリンとしての公共政策学を主張する最も有力な論者は足立幸男であろう。足立は、日本公共政策学会の2003年度研究大会における会長基調講演およびその論旨を敷衍した著作の中で、公共政策学が独自のディシプリンであること、公共政策の研究や実務にはポリシーマインド・政策学的思考という思考ないし能力が必要となることを主張している（足立編著 2005および足立 2005）。

足立によれば、ポリシーマインドとは、システム思考と社会構成員の公共政策の決定当事者としての自覚、およびその自覚に基づいた公共政策デザイン力の獲得努力である。

システム思考とは、ある種のバランス感覚をもって社会構成員全体の集合的利益を最大化するよう公共政策の分析・評価やデザインを行うことであり、一方では公共政策システムないし政策体系の全体を視野に入れ、その全体のバランスを勘案しつつ個別の政策・施策をデザインすることを求め、他方では個別の政策・施策の実態からシステム・体系の全体を認識することを求める思考様式である。また、福祉や環境、防衛といった特定の政策分野や自由や平等、安全といった特定の公共価値を絶対視しないことや、実施費用の可能な限りの節約を含めた機会費用への最大限の配慮を求めるなどポリシーマインドの内容は幅広く、その外延は可動の余地がありそうである。

政策決定者としての社会構成員の自覚というが、強調されるのは社会構成員の大部分を占める普通の人々としての市民の役割である。それは、社会に発生する様々な問題のうち、どれが公共的課題なのかということは客観的に存在し発見されるものではなく、主観的に構成され、社会構成員の間の政治を通じて決定されるという認識に基づく。また、政治家や政府職員の行動、政策デザイン、政策評価等に影響を与える倫理や公共哲学も同様に構築される（足立・森脇編著 2003、pp. 13-14）。そこで、このように公共政策に対し、個々人としてはともかく集合的には大きな影響を及ぼす市民が、バランス感覚を持ってよりよい社会を構築するという方向で思考し行動することが重要となる。そのために求められるのが、政策決定の責任者の観点に仮想的に我が身を立たせて政策デザインや決定、評価のよしあしについて考えることであると足立は論じる。

上記のようなバランス感覚や自覚を具体的な政策・施策の形にするのがデザイン力である。首相、大臣、国会議員、官僚、マスコミ、利益集団、市民など、アクター毎に必要とされる能力は異なってくると思われるため、政策デザインのガイドラインを構想するなら、本来、アクター毎に異なるものが必要となると思われるが、デザイン力が必

要であるには違いない。こうして、市民を含めた社会構成員が統治の主体という自覚を持ち、システム思考に基づいて政策デザインを行うなら、社会の各層や各方面から政策提言や提言に対する議論が発信され、アイディアの競争（competition of ideas）が発生し、様々な政策提言の中からよりよい政策が採用され、よくない政策は淘汰されることが期待される。

このアイディアの競争という概念は、社会的問題解決の戦略であるインクリメンタリズムで知られるアメリカの政治学者・政策学者のC.E.リンドブロム（Lindblom, C. E.）が提示したものである。リンドブロムが追求するのは、社会のさまざまな公共的課題が解決され、よりよい社会が実現することであり、そのために不可欠とされるのが、民主的かつ合理的な政策形成である。合理的であるためには、アイディアの競争が機能しなくてはならない。そのためには社会が民主的であること、すなわち社会の構成員の政治的影響力がおおむね等しく、突出した政治的影響力を行使するアクターが存在しないという状態を実現する必要があるというのがリンドブロムの主張の要点である。インクリメンタリズムを提唱した *A Strategy of Decision*（1963）やそれに続く *The Intelligence of Democracy*（1965）といった著作では、こうしたプロセスが規範であるとともに、アメリカのような先進民主主義国では既に実現しており、あたかも市場における競争を経て最適の商品やサービスが勝ち残って社会を発展させているように、アイディアの競合を通じて最適の政策が勝ち残って社会を発展させていると主張し、当時アメリカ社会がベトナム戦争等で行き詰まりを感じていたこともあって厳しく批判された。その後のリンドブロムの研究生生活は、アイディアの競合を目指すべき目的と捉え直し、民主的かつ合理的な政策形成を阻む原因の解明とその実現方策の探求に捧げられた。その成果が最も接近しやすい形でまとめられたのが1968年（初版）、1980（第2版）、1993（第3版）と版を重ねた *The Policy-Making Process* である。リンドブロムの研究の進展にともないその内容は版によって変化しているが、とりわけE. J. ウッドハウス（Woodhouse, Edward J.）を共著者に迎えた第3版では大きく変化し、政策過程にお

ける経済界の影響力が不当に大きく、社会の問題解決能力を減損させ、公共政策をゆがめているという従来からの主張を一層強調するとともに、そうした状況を改善する処方箋を示している。

リンドブロムらの提示する処方箋は、公共政策についての研究の役割の見直しと経済界の特権的地位の抑制であり、その目指すところは、社会構成員の影響力や社会構成員が公共政策や社会の将来像をデザインする環境や能力の平準化である（Lindblom & Woodhouse 1994, p. viii）。とりわけ、公共政策の分析や評価—政策形成に関わる事実収集、研究、議論、調査、情報処理—の役割を、政治的エリートに情報を与えるのではなく、一般の人々が社会問題と社会が持つ可能性について、より明瞭に考えられるよう支援することとしている点は重要である。リンドブロムらがこのように主張する根拠は、分析・評価は無謬ではあり得ない、自由と安全・安心のトレード・オフといった価値の対立に唯一の正解を出すことはできない、政策決定に必要とされるタイミングや実行可能性の観点から許容される費用の範囲内で正確かつ有用な情報^{インテリジェンス}を提供することは困難、公共的課題は客観的実在ではなく社会問題の中から主観的に構成されるものであるがゆえに分析・評価ではどの問題に取り組むべきかを確定的に示すことはできないといった限界が公共政策の分析・評価にはあるという事実である（Lindblom & Woodhouse 1994, 第2章）。

足立やリンドブロムの議論は、一般の人々としての市民の役割を重視し、政策分析や評価なども市民への情報提供手段であり、比較的中長期の漸進的な改革を意図しているといえる。本稿もこの議論に依拠するものであるが、これらの議論が普通の人々としての市民の役割を強調するとしても、それは市民だけが重要だというのではなく、むしろ政府職員など市民以外の社会構成員の役割、特に政治的指導者のリーダーシップが重要であることを自明とした上で、従来見過ごされてきた市民の役割を強調している議論だということには注意を要する。また、リンドブロムとウッドハウスが主張する経済界の特権的地位の抑制については、有効かつ弊害の少ない具体策をデザインすることは困難であり、むしろ可能な限り経済界以

外のアクターの地位を向上させることで平準化を図るべきと考えられる。

公共政策学が独自のディシプリンであるとして、そこにはどのような研究領域があるのだろうか。大別して3つの、詳細には11の研究領域があるとされる（足立編著 2005、pp. 3-6）。この議論については足立による先駆的な業績があるだけであり、より一層の活性化が期待される。

公共政策学が独自のディシプリンであるとして、以下のような研究領域があるとされる（足立編著 2005、pp. 3-6）。

①世界観や政治文化の研究。自由や経済発展など公共政策のよしあしを判断する規準となる諸価値の間の重み付けが、ある社会でどのようなになっているかその実態を解明する研究。②公共政策のデザイン・決定・実施・終了の実態の解明。③政策形成に関わる制度と実際の公共政策の因果関係の解明やよりよい公共政策が形成される可能性の高い制度の構想。④政策形成過程のアクターの研究。思考や行動、相互影響関係の実態解明やそれらをより公共的なものに代えるための方法の研究。これらは公共政策についての理解を深めようとする研究といえる。

⑤政策過程の研究。ある特定の政策のデザインから決定、実施、終了に至るプロセスについての情報を収集・分析し、政策改善のための教訓を得ようとする事例研究。失敗した政策をケースとして取り上げる場合は、いわば公共政策の「失敗学」になると思われる。⑥実務としての政策分析・政策評価。個別具体的な公共政策・施策を分析・評価しようとする研究。⑦個別具体的な公共政策のデザインと提言。これらは実務としての公共政策研究とされる。

以下の4つの研究領域が個別分野の政策実務への寄与を目的とした公共政策の一般理論・モデル・手法の研究であり公共政策学原論を構成するとされる研究領域である。⑧政策類型の研究。政策手法（policy instrument）にどのようなものがあるか、各手法はどのような特徴を有するかを整理する研究。⑨政策分析、政策評価、政策デザインのための一般理論および手法の研究。⑩ポリシーマインドの研究。⑪学説史の研究。

やや引用が長くなったが、ここでは政策評価論

が取り組んできた政策評価のための一般理論および手法の研究が公共政策学の原論として重要なポジションを与えられ期待を寄せられていることに注目すべきであろう。政策過程の各段階において、社会調査等の手法を応用した評価手法を活用し、公共政策をよりよい方向に統制し、よりよい公共政策や政府を実現しようとするのが政策評価論であり、そうであるならば公共政策学との親和性は高く、その中核にふさわしいといえる。

では、公共政策学の研究領域としての政策評価の一般理論にはどのような課題があるのだろうか。以下の3つの課題があると考えられる。第1に、評価が物事の価値を判じ定めることを意味する以上、よい政策とはいかなるものか、よくない政策とはいかなるものかを明らかにすることが求められる。第2に、よい政策のつくりかた・政策デザインのガイドラインの探求が課題となる。第3には、よくない政策を改善または終了することが求められる。政策実施論や政策終了論が明らかにしてきたように、政策の改善や終了は容易ではない。そこで、既存の政策の継続・改廃を決定し実行する技術や手法の開発とそのために導入されたシステムや制度の批評と改善方策の提言が政策評価論の課題となる（窪田 2004、p. 37）。これらの課題がバランスよく探求されることが政策評価論の健全な発展のために必要であるが、現状を見れば、第3の課題、とりわけ国や地方自治体が導入した評価制度の批評・改善案の提言に終始している点がわが国の政策評価論の特徴であり、問題視せざるを得ない点である。何らかの改善展開が求められよう。これについて具体的に論じる前に、政策評価論の展開とその中で日本型政策評価の位置づけを明らかにする必要がある。

3. 政策評価論の展開

政策評価論という研究領域は、アメリカ合衆国における理論研究と実践の相互作用の成果によってリードされてきた。その展開について、評価対象となった政策がよかったかよくなかったかの判定を重視するタイプの評価から改善策のデザインを重視するタイプの評価に進化したと整理するこ

とも可能である（窪田 2005、第1章）。しかし本稿では、当初、実施中の政策のインパクトを事中・事後に科学的に妥当な方法で正確に測定し、分析、評価することからスタートした政策評価が、事前、すなわち政策デザインの段階を含む政策形成の全段階にまで評価の対象を拡大したという展開に注目したい。

政策形成過程や政治・行政の実務、軍事等への社会調査手法の活用という意味での評価の起源は、1930年代に求められるとされる。しかし、今日的な意味での政策評価の起源は1960年代の後半、GAOによるプログラム評価の導入に求めることができる。理論面では、科学的評価を主張するD. T. キャンベル（Campbell, Donald T.）と実践的评价を主張する L. J. クロンバック（Cronbach, L. J.）らとの論争が重要である（Campbell 1969およびCronbach 1980）。キャンベルが、実験計画法のような手法を使用して科学的に妥当な結果を出すことができた場合にのみ、その結果を公共政策の継続・廃止の決定に活用すべきとしたのに対し、クロンバックらは政策決定に適切なタイミングで、政策決定に有用な情報^{インフォメーション}を提供することこそ政策評価の役割であるとした。

政策評価に科学（science）と技・術（art）という2つの側面があるとして、キャンベルは科学の領域に留まり、その範囲内で実践に貢献することを主張したのに対し、クロンバックらは科学を基礎としつつも技・術の領域に踏み込んで実践に貢献する、あるいは実践を改善することを主張したわけである。もっとも、実践的评价といっても、真実ではない情報を提供しても意味はない。そこで、この論争は、政策過程における現実的制約の中で、情報の正確さをどこまで妥協するかというとらえ方をすべきであろう。キャンベルは妥協を排し、クロンバックはある程度の妥協を許容するということである。

キャンベルの主張に従えば、公共政策に関して社会学者が失敗することはないだろう。しかし、社会科学の潜在的能力・可能性を十分に活かすことにはならないし、社会学者が公共政策により積極的に関与する場合と比べて、公共政策が失敗する可能性は高い。そこで政策評価論の研究者の多くは実践的评价を支持する立場をとり、社会科

学者が結果的に失敗するかも知れないという危険を認識しつつ、公共政策に関与すべきであると考ええる。研究者の多くが実践的评价派である以上、科学的評価派と対比してひとくくりにすることは無理があり、さまざまなアプローチが存在する。個別の研究者とそのアプローチを紹介することは本旨からはずれるため、ここでは少なくとも、①どこまでの妥協を許容し、正確さを犠牲にするかという軸と、②誰に情報を提供するのか、政府の舵取りを行うトップリーダーなのか、個別の政策や施策の責任者（manager）なのか、政策や施策の対象者や利害関係者なのか、広く社会構成員なのかという軸が存在することを指摘するにとどめたい⁴。

公共政策に関与し、実践で活用されることを意図する以上、実施中の政策のインパクトを事中・事後に科学的に妥当な方法で正確に測定し、分析、評価するインパクト評価から出発した政策評価が、事前、すなわち政策デザインの段階を含む政策形成の全段階にまで評価の対象を拡大していったことは必然であった。公共政策のデザインや決定、実施に関わる諸アクターにとって、当初の政策評価が提供しようとした、個別の政策・施策が成功したか失敗したか、あるいはより狭く所期の目的を達成したかできなかったかという情報だけでは不十分であり有用性が乏しかったからである。また、政策評価の研究者と実務家からしても、公共政策と政府を改善し、よりよい社会を実現するという意図からして、事後的に、許された時間や資金の制約の中で、なるべく正確に政策・施策の実施に起因する変化、すなわちインパクトを測定、分析し、所期の目的を達成していない場合はそれを理由として廃止や改善を勧告するだけでは不十分であったからである⁵。政策評価論において、通常、政策評価が、政策形成過程の各段階における体系的な社会調査活動と定義されるように⁶、そもそも目的が適切に設定されているか、プログラム理論は適切か、実施過程は政策に従って正確であり効率的かといった点もまた評価により明らかにされ、問題があれば改善されるべきと考えられたからである。

政策評価とは、政策過程の開始から終了に至る全段階で、それぞれの段階において明らかにすべ

き事項に即した社会調査手法を活用し、よりよい政策づくりや、正確かつ効率的な実施や、よくない政策の継続休廃に貢献し、もってよりよい社会の実現を図ることであるという主張の代表的な論者はロッシらである。ロッシらは自らの政策評価を体系的評価 (systematic evaluation) と名付けているが、それが政策過程の全段階に関与するものであることから全段階評価 (all stage evaluation) と名付けることも可能である。

ロッシらが体系的評価の名のもとで行うべきとするのは、社会問題が発生した段階でニーズの有無・程度を明らかにするニーズ評価 (needs assessment) であり、政策デザインの段階でプログラム理論を妥当なものとするとともに事中・事後の評価が可能となるよう調整する評価可能性評価 (evaluability assessment) であり、プログラム理論が正確に実行されているか、あるいは効率性において改善すべき点はないかを明らかにするプロセス評価 (process monitoring) であり、政策・施策に起因する変化を正確に測定するとともに成功または失敗の原因を分析し明らかにするインパクト評価 (impact assessment) であり、政策・施策の効率性を明らかにするための費用便益分析 (cost-benefit analysis) や費用有効性分析 (cost-effectiveness analysis) である。もちろん、これら全てをロッシらが開発したわけではない。費用便益分析や費用有効性分析は経済学で開発された手法であるし、政策評価を事前段階に拡張したものとして重要な評価可能性評価は、J. S. ホーリー (Wholey, J. S.) が開発したものである。ロッシらの貢献は、さまざまな手法を政策過程の各段階に適切に配置し体系的に実行することを提唱した点にある。ロッシらのテキストである *Evaluation: A Systematic Approach* が版を重ねつつもアウトラインは変更されず、しかも支持され続けている点が証左となろう (Rossi, Lipsey, Freeman 2004)。

ここで注意すべきことは、政策評価というものの、単一の手法が政策形成の各段階で活用されるわけではなく、社会調査や経済学など、さまざまな分野で開発されてきたさまざまな手法が用途に応じて政策形成の各段階で活用されることである。よって、政策評価が政策形成の各段階に拡張されたというより、政策形成の各段階で活用可能な社

会調査等の手法を政策や政府の改善、ひいてはよりよい社会の実現に活用することをさして政策評価というと解した方が正しい。いずれにせよ、体系的評価では、まず、ある社会問題に対して政府が対処することにニーズがあるかを明らかにし、ニーズがある場合には、適切なプログラム理論が構築されるよう支援し、政策決定後に実施が開始されたらプログラム理論に従って正確に実施されているか、効率的に実施されているかを監視・モニタリングし、政策・施策に起因する変化を正確に測定・分析し、その結果に基づいて政策・施策の継続休廃を勧告するとともに改善を勧告する場合は具体的な改善策をデザインし、何らかの必要がある場合には費用便益分析や費用有効性分析、あるいはその他の手法を用いて他の政策・施策との相対順位を明確化する。問題となるのは、このような政策評価を全ての政策・施策に対して行えばよいのであるが、そうすることは人材や費用の面から実行可能性を欠くため、どの個別的な政策・施策に対してこのような体系的評価を行い、どれに対しては行わないかであるが、定説はないのが現状である。

なお、最近各国で導入が進み、注目されている業績測定 (performance measurement) については、その重要性についてはほとんどの研究者が否定しないものの、既存の理論と同一の体系に整合される所までは消化されていない。しかしながら、政策・施策の有効性の有無の判断を正確さを犠牲としつつ簡易に判断する手法であり、より複雑で高価なインパクト評価を行うべき政策・施策を見極めることに有用な手法と認識されているか、成果主義に基づいて公共政策の実施にあたる行政機関等を統制する手法と認識されているかのどちらかである。

4. 日本型政策評価とその課題

わが国の政策評価の実践を見れば、特に地方自治体において、その多くは三重県が導入した事務事業評価をベースとしてその発展上にある。諸外国で導入が進んでいる業績測定と事務事業評価を同一視すべきではない。事務事業評価では、業績

測定よりはるかに多い事項について情報が収集され、判定が行われる。手法という点で両者は異なるわけであるが、その他にも発想や使用法の点で、事務事業評価は業績測定とは異なる。わが国の地方自治体発祥のモデルであり、他に類例を見いだしにくいことから日本型政策評価と名付けることができる。

日本型政策評価としての事務事業評価については、プログラム評価と業績測定を混淆、折衷案として行われていることが多く、いずれの長所も出ないまま役に立たない評価として行われていると評される（山谷 2007、p. 8）。首肯できる部分の多い批評であるが、事務事業評価については、もう少しポジティブに評価してもよいのではないかなと思われる。

事務事業評価は、日本版の業績測定というよりもむしろ、ロッシらが提唱する体系的評価を、安価かつ政策体系の全体に網羅的に行える自己評価として行うことを目指したモデルであると捉えるべきである。それは事務事業評価で作成される評価表に含まれる項目に示されている⁷。

事務事業評価では、政策体系の各レベルの担当者が、自分の担当する施策や基本事務事業、事務事業について、上位の政策とどのような連関があるか、目的としてどのような状態を実現しようとしているのか、そのために誰が、誰または何に対して、どのような行為を行うのかといったことを明らかにするとともに、その実現状況を明らかにする成果指標を設定する。精緻さには欠けており、評価表を作成する行政職員の無理解や能力の限界から運用にも問題があるものの、その目指すところはセオリー評価と同じである。

事務事業評価において、ある事務事業は事前および実施中に評価される。終了時やそれ以降という意味で事後的に評価されることはないが、実施中は毎年度1回、定期的に評価される。事前評価用と事中評価用にそれぞれ別の評価表を用意している自治体もあるし、同じ評価表を用いている自治体もある。いずれにせよ、事前評価に用いられる評価表には、自治体により差違はあるものの、ニーズの有無や強度を問う項目が含まれている場合が多い。行政職員が、特に社会調査を行うことなく、業務上で得られた感触に基づいて記入する

ため、正確さは低いレベルに留まるが、ニーズ評価に相当するものであるといえる。また、公共政策としての妥当性を判断する項目も通常に含まれている。事務事業が解決しようとする問題が公共的課題といえるか、公共的課題であるとして国が都道府県か市町村かNPO等か、いずれが対処すべきかが判断される。市場化テストという形で政府の業務にこの種の判断を持ち込むことは、イギリスやニュージーランドが先鞭を付けているが、個別の事務事業の採否の判断について行うのは事務事業評価の独自の特徴である。この点からも事務事業評価が業績測定よりもむしろ体系的評価に近いものであることは明らかである。

さらに、事例は報告されていないが、事務事業評価表を活用すれば、プロセス評価を行い、計画と実施の齟齬を明らかにし、修正することも理論上は可能である。

また、インパクト評価との関連では、事務事業評価は政策目的の実現状況を数値で正確に測定しようとすることはもちろん、副次的効果や弊害についても、実施担当の行政職員が業務中に認識できた限りにおいてではあるが、明らかにし、継改休廃の判断材料として用いようとしている。もっとも、副次的効果や弊害を視野に収めているのは、三重県の事務事業評価システム以外にそう多くの例はないかもしれない。

費用便益分析や費用有効性分析が事務事業評価にそのままの形で取り入れられることはなかった。自治体職員の能力や投入可能なコストの観点から不可能だったのであろう。しかしそれらが目指すところの政策間の順位付けは、異なる形ではあるが、事務事業評価においても追求された。三重県で一時採用された劣後表 一部局単位でそれぞれが担当する事務事業を優先順位の低い順に並べるリストや注力表示 一ある施策の中でそれを構成する事務事業がどの程度重要であるか、今後どの程度重視すべきかを3段階程度で示すもの のように直観に基づいて順位を定める単純な方法もあれば、評価表に含まれる複数の項目に得点を割り振って、獲得点に応じて事務事業の順位を明らかにしようという方法もあった。いずれも費用便益分析や費用有効性分析と比較すれば、客観性において劣ることは明白であるが、何らかの

手法で政策間の相対順位を判定し、予算編成等の政策決定に反映させることが必要であるという認識、あるいは反映させたいという志向が存在したことの証拠とはいえる。

もちろん、三重県をはじめとする事務事業評価を導入した地方自治体が、事務事業評価と業績測定を同等視する傾向も見られた。しかしながら、成果主義の人事や組織管理といった、一般に業績測定に期待される役割については、活用を期待されることはあっても事務事業評価が実際に担うことは少なく、その面で確たる成果を上げることもなかった。

やはり事務事業評価は、自己評価によって、安価かつ網羅的に体系的評価を導入しようとした試みと認識されるべきであろう。しかし、残念なことに、実行可能性への配慮に欠けており、一時的な改革努力としてはともかく、このままの形での定着は困難と思われる。その理由は、端的に言って、自治体職員の能力や負担可能なコストを無視しているという意味で単純に実行可能性を欠いているからである。社会調査や経済統計や会計の知識や手法を活用して政策の効果やコストを正確に測定分析することは、現在の自治体職員が受けている採用試験や採用後の研修からは困難であろうし、能力があったとしても、政策の実施に関わる業務をこなしつつ調査・分析を行う時間的余裕や調査・分析に必要な予算を捻出することは不可能である。さらに、リンドブロンらが指摘するように継続・改善の判断に必要なタイミングで必要な評価結果・情報^{インテリジェンス}を提供することは困難である。最近注目されている外部有識者による評価委員会が行う外部評価を導入しても、評価委員会が独自の調査を行わないなら、自己評価と同じように正確な評価を行うことはできない。ここで確認すれば、体系的評価と呼ばれる政策形成の各段階における社会調査手法の活用を行い、予算編成などの意思決定に必要とされるレベルの正確さを実現するためには、それ相応のコストや人材が必要であり、さらにいえば、必要とされる正確さを保ちつつ、適切なタイミングで結果をまとめることは至難である。

このように、事務事業評価は日本型政策評価と呼ぶに足る独自の発想の産物であり、体系的評価

という最新の政策評価理論に沿ったものであって、その点は高く評価されてもよいが、実行可能性に難がある。政策評価論で論じられる体系的評価では、あまり強調されないが、評価対象を全政策のごく一部に限定し、選択的に評価を行うという形でこの問題に対処している。にもかかわらずわが国で自己評価かつ基本的に全数評価として導入されたのは、わが国の地方自治体の職員の文化において横並び式の一斉導入でないと受け入れられなかったからという事情があると説明される⁸。また、政治的リーダーや自治体職員が、公共政策学や政策評価論の学知に理解を欠いていたことも指摘せざるを得ない。知識があればこのような困難な挑戦はなされなかったのではないと思われる。

こうした事務事業評価の導入であるが、国の府省の政策評価制度をはじめとする評価関連の制度の導入の契機となったことは貢献として記録されるべきであろう。しかしながら、既に見たように実行可能性に問題があったため、そして事務事業評価を何に使用するかという導入目的に問題があったために適切に発展することができていない。

財政改善のための事業カットに事務事業評価を導入するというのは誤った導入目的の典型である。確かに事務事業評価を最初に導入した三重県の公式文書にも、他の目的と並んで財政への貢献が掲げられていた⁹。その背後にあり、実務で強調されていたのは、必要性や有効性、費用などに問題があるのに継続されてきた事務事業を終了させ、浮かせた人員や予算をより必要性の高い問題や、有効性の高い事務事業に転用するという発想であった。ところが、多くの地方自治体では、単純に歳出の金額の削減を目指し、事務事業評価を導入した。歳出削減のためには、ニーズ等の必要性や有効性が一定あり、費用の面でも大きな問題がないような事務事業であってもカットせざるを得ない場合が多い。事務事業評価は一般に正確さに劣るが、その中でも事務事業間の相対比較については特に不正確な結果しか出せない。そしてその割には手間暇がかかる。よって、歳出削減のためには、事務事業評価とは別に、期限と目標金額を定めて集中的に事務事業の見直しを図った方がよい¹⁰。多少なりとも必要性があり、有効性もあ

る事務事業をカットするのであるから、政治的抵抗は大きいと覚悟せざるを得ない。その点を踏まえて、財政部門主導で行うか首長直轄の特命チームで行うか、後者の場合、チームを職員で構成するか、外部有識者で構成するか、住民を入れるかなどの選択肢から、状況を見て適切なものを選ぶしかない。事務事業評価を導入すれば歳出額が減ると期待したのにそうならなかったと不満を述べる実務家も多いし、それが今日の事務事業評価の停滞につながっている面もあるが、最初から実現するはずのない期待であったと言い過ぎであらうか。

アカウンタビリティの確保もまた、多くの自治体が事務事業評価に期待した導入目的である。しかし残念ながらアカウンタビリティのジレンマという、アカウンタビリティ確保のための作業負担が増え続け、物理的には本来業務にまわすべき労力を大きく浸食し、本来業務の実績（パフォーマンス）が低下して成果が出ないという状況に陥っている（山谷 2006、p. 210）¹¹。

まして、事務事業評価を人事評価に活用したいという期待は、事務事業評価のなんたるかを理解していない議論といわざるを得ない。事務事業評価は、毎年の予算編成等の政策決定に間に合うようなタイミングで、事務事業の採否・継続改廃の判断に必要とされる事項を、自己評価で可能な範囲と正確さで収集・分析したものである。そのような目的も異なれば、正確さも異なる情報に基づいて人事が決定されとなれば、士気と意欲の低下は不可避である。それだけでなく、保身のために評価がゆがめられるという弊害が発生する。

事務事業評価の今後としては、自己評価でこそ正確に明らかにできる事項に特化すべきである。それは、ロッシらの用語を用いれば、セオリー評価でありプロセス評価である。これらは、政策を担当する政府職員こそが正確に行うことができる¹²。

政府が全ての政策・施策についてセオリー評価を行うということは、何が公共政策によって解決されるべき問題なのか、目的としてどのような状態を実現しようとしているのか、そのために誰が、誰または何に対して、どのような行為を行うのか、それらは上位の政策を含めて他の政策とどのような連関があるか、といったことを真剣に考慮する

とともに、社会に公表するということである。そうしたことが行われることは、公共政策学が必要と強調する一般市民が社会問題や政策選択肢を考える環境づくりや考える能力向上に資するところ大である。事務事業評価に果たし得る役割は、公共政策のデザインの活性化である。事務事業評価によって、政府のどのような部局が、どのような政策・施策を行おうとしているのか、行っているのかを知るためのコストは激減する¹³。わずかな時間や手間ですうした情報が得られるなら、よりよい公共政策をデザインし提言しようというアクターは増えるだろう¹⁴。同じく全ての政策についてセオリー評価を踏まえてプロセス評価（モニタリング）を行うことはアカウンタビリティの向上に資するところ大である。また、一般市民であれ、専門家であれ、インパクト評価や費用便益分析等を行うことが容易になる。

インパクト評価や費用便益分析等については、自己評価で正確な評価結果を出すことは至難であるから、それが必要とされる一部の政策・施策について外注によって、またはそれ専門の政府機関を設置してそこで行うことを基本とする。費用便益分析等の政策・施策間の相対順位を明らかにする評価手法については正確さという点でインパクト評価に劣るので、公共事業の事前評価など特に必要がある場合は別として、一般にインパクト評価を優先させるべきであろう。

5. 政策評価と公共政策のデザイン

本稿では、政策評価論において、体系的評価という評価理論が有力となっていること、および日本型政策評価としての事務事業評価も、多分に無意識ながらもその考えに沿ったものであることを確認した。

政策評価が体系的評価として政策形成の各段階の全てで行われる評価に進化しようとするなら、それは政策デザインという価値判断を不可避に迫られるとともに科学（science）と技・術（art）の両方が必要とされる領域に踏み込むことであり、政策形成に対する市民の意欲を喚起し能力を高めることが必要である。この点について政策評価の

理論と実践があまりに無自覚であることが重要な問題であると思われる。

ある公共的課題に政府が対処すべきであるか、優先順位はどの程度か、どのような政策目的を設定すべきかといった問いに答えるためにはニーズ評価を行うべしと体系的評価は主張する。しかし、ニーズの有無・大小は判断材料の一つではあっても全てではない。顕在化したニーズを充足することが公共政策の課題というのはあまりに矮小化された視点である。将来世代を含む社会のマイノリティに関わる政策・施策が代表的であるが、顕在化したニーズが小さい、あるいは無くても時として公共政策が必要となる場合はある。また、ニーズが大きい場合も、世論が求める政策手法とは別の手法を世論に反してでもとるべき場合もある。

複数の解決策・代替案を作成し、費用便益分析や費用有効性分析を行って最善案を選択せよと主張する論者も多い。疑いなくそうすべき場合もあるが、全ての公共政策についてそうせよというのは実行可能性を欠く。また、そもそも多種多様に存在する政策手法の中からどれを選択すべきか、適切な政策手法をどうやって発想するかという解決策・代替案作成にあたっての重要問題であり、科学を基本としつつも属人的な技や術を不可避に必要とする問題について、体系的評価は完全に沈黙している。

また、事務事業評価に限って言えば、よい事務事業の集積がよい公共政策とは必ずしもいえないという事実を自覚することが必要となる。事務事業評価を行い、よい新規事務事業を増やし、実施中のよくない事務事業を改休廃することには大きな意義があるが、よい公共政策の体系を構築するには、そうしたボトムアップのアプローチだけでなく、政策から施策、施策から事務事業をデザインするトップダウンのアプローチが必要とされる。そのためには、少なくとも政治的なリーダーシップが必要とされるが、多くの地方自治体にはそれが欠けていたように思われる。

公共政策や政府をよりよい方向に統制し、それによってよりよい社会を構築しようとするならば、体系的評価が主張するように、公共政策の形成過程の各段階で社会調査等の手法を用いた予測・測定・分析・評価を行い、その結果を活用す

ることも確かに必要ではあるが、同時にそれだけでは不十分であり、あるべき社会や人間の姿を描く公共哲学と公共政策や政府がそれに従い踏み出すべきではない倫理を考察・構築し、それを踏まえて評価することもまた必要である。特に、既存の政策評価論がそうであったように、特定の政策・施策をよりよい方向に統制することに関心を集中させるだけでは不十分であり、公共政策システム全体をよりよい方向に導くという発想を重視する必要がある。

こうした点について研究を行い、実践に反映させることが政策評価論の課題である。図1に即していえば、理論から実践に向かう作用を拡充させるということである。そして、これらを行うことにより、政策評価は興隆しつつある学問分野である公共政策学の中核に位置づけられるだろう。公共政策学も政策評価論もよりよい公共政策によってよりよい社会を構築しようという目的は共通だからである。具体的な産出物としては、公共政策のデザインから決定、実施、終了を一貫して網羅し、それに従えばたとえ素人であっても一定レベルの公共政策をデザインできるとともに、他者がデザインした公共政策の是非を判断できるようなガイドラインを構築することが政策評価論に求められると考えられる。

なお、多種多様な公共政策・施策のうち、どれに対し政府がインパクト評価を行うかという問題は、非常に難しい問題である。議会の決議、首相または首相の諮問機関が指定する、学識経験者による委員会を設置しその勧告によるといったアイディアが浮かぶが、詳細は今後の研究課題である。また、政府が行うインパクト評価や費用便益分析等の外注の受け皿となる大学やシンクタンクやNPO等の充実もまた緊急の課題であろう。

謝辞

本稿は、第8回全国大会発表要旨集録に掲載された原稿「日本型政策評価とは何か」に大幅に加筆し修正を加えたものである。大会における報告に、貴重なご指摘をいただいた方々および匿名の査読者の方々に記してお礼申し上げたい。

注記

- 1 政策を狭義に解せば、事務事業評価の多くは政策レベルの評価ではないため、事務事業評価は政策評価ではない。しかし、政策を広義に解し、公共問題解決の方針とその処方箋の双方を指すとすれば、処方箋すなわち事務事業もまた政策であり、それを対象とする事務事業評価もまた政策評価の一種である。
- 2 政策評価における理論と実践の相互作用については、E.R.ハウス（House, E.R.）とK.R.ハウ（Howe, K.R.）*Values in Evaluation and Social Research* を参照のこと（House & Howe, 1999：p.x）。
- 3 企業の政策には、自らの経営のためのいうなれば企業の私政策だけではなく、企業による公共政策も含まれる。
- 4 政策評価論の展開を、主要な論者に着目して解説したものととしてW.R.シャディッシュらの著書 *Foundations of Program Evaluation: Theories of Practice* が有用である（Shadish, Jr., Cook, Leviton, 1991）。
- 5 評価対象の政策・施策を廃止すべきか否かの判断を重視するのが判決型評価、廃止せずに改善するとして、どのように改善すべきか、それはなぜかを重視するのが分析型評価である（窪田、2005：第1章）。
- 6 代表例としては、P. H. ロッシ（Rossi, P. H.）らによる定義（Rossi, Lipsey, Freeman, 2004：p.3）を参照のこと。
- 7 三重県の事務事業評価システムで使用された評価表については年度や用途の異なる各種タイプが存在する。拙著を参照のこと（窪田、2005：参考資料）。
- 8 筆者が事務事業評価を担当する三重県の幹部職員に面接調査した際の聞き取りによる。
- 9 長期計画策定・評価システムと事務事業評価システムと予算編成システム・執行管理システムが相互に関連し、それらが中期財政政策と関連するとされた（窪田、2005：p.62）。
- 10 三重県でも1998年度～2000年にかけて公的関与による見直しやスプリングレビューと呼ばれる事務事業見直しを事務事業評価と平行して実施し、一定の成果を上げている（窪田、2005：p.92）。
- 11 ここでも正確な評価を行うことが必要となる。外部評価にあたる有識者の権威でアカウンタビリティを確保しようとするのは間違いである。
- 12 むろん、研修等による適切な訓練が必要であること

はいうまでもない。

13 この点から、インターネット上で個別の評価表を公表されることが必要不可欠である。

14 また、副次的には、公共政策や行政についての研究の促進が期待できる。

参考文献

- Campbell, Donald T. (1969). Reforms as Experiments. *American Psychologist*, 24
- Cronbach, L. J., Ambron, S. R., Dornbusch, S. M., Hess, R. D., Hornik, R. C., Phillips, D. C., Walker, D. F., & Weiner, S. S. (1980). *Toward Reform of Program Evaluation*. Jossey-Bass.
- House, Ernest R., Howe, Kenneth R. (1999). *Values in Evaluation and Social Research*. Sage.
- Lindblom, Charles E., & Woodhouse Edward J. (1993). *The Policy-Making Process (Third Edition)*. Prentice Hall. (チャールズ・E・リンドブロン／エドワード・J・ウッドハウス著、藪野祐三／案浦明子訳（2004）『政策形成の過程：民主主義と公共性』、東京大学出版会）
- Rossi, Peter H., Lipsey, Mark W., Freeman, Howard E. (2004). *Evaluation: A Systematic Approach (Seventh Edition)*. Sage. (ピーター・H・ロッシ／マーク・W・リプセイ／ハワード・E・フリーマン著、大島巖／平岡公一／森俊夫／元永拓郎監訳（2005）『プログラム評価の理論と方法：システマティックな対人サービス・政策評価の実践ガイド』、日本評論社）
- Shadish, Jr., William R., Cook, Thomas D., Leviton, Laura C. (1991). *Foundations of Program Evaluation: Theories of Practice*. Sage.
- 足立幸男・森脇俊雅編著（2003）『公共政策学』、ミネルヴァ書房
- 足立幸男編著（2005）『政策学的思考とは何か：公共政策学原論の試み』、勁草書房
- 足立幸男（2005）「政策研究：規範、倫理、公共性」、北川正恭・縣公一郎／総合研究開発機構編『政策研究のメソドロジー：戦略と実践』、法律文化社
- 窪田好男（2004）「政策評価の研究」、『人環フォーラム』第15号、京都大学大学院人間・環境学研究科
- 窪田好男（2005）『日本型政策評価としての事務事業評価』、日本評論社

政策分析ネットワーク編（2003）『政策学入門：ポリ
シースクールの挑戦』、東洋経済新報社
山谷清志（2006）『政策評価の実践とその課題：アカ
ウンタビリティのジレンマ』、萌書房
山谷清志（2007）『政策評価の質とその改善－実践と

研究の交錯』、『評価クォーターリー』第1号、（財）
行政管理研究センター
龍慶昭・佐々木亮、（2004）『「政策評価」の理論と技
法：増補改訂版』、多賀出版

（2008.1.23受理）

Public Policy Studies, Evaluation Studies, and the Mode of Evaluation in Japan

Yoshio Kubota

Associate professor of The Faculty of Law, Kobe Gakuin University
kubota@law.kobegakuin.ac.jp

Abstract

Public policy evaluation is expected to be a central issue of public policy studies which is now emerging as a discipline. But the reality of introducing systems of evaluation and the arguments surrounding those are disappointing.

One reason is that local governments that introduce a Japanese style policy evaluation system misuse it. Another reason is that Japanese policy evaluation studies concentrate on the critique of introducing a Japanese-style public policy evaluation system.

The role of the Japanese style public policy evaluation system needs to be re-defined to activate competition of ideas.

Keywords

public policy studies, policy evaluation system in Japanese local government,
topics of policy evaluation studies, purpose of policy evaluation, public policy design

【実践・調査報告：依頼原稿】

自治体評価と外部支援指導

星野 芳昭

経営改革プロデューサー、JMAC構造改革推進セクター
hoshino@stones.com / yoshiaki_hoshino@jmac.co.jp

要 約

本論文は、自治体が行政評価制度を導入して定着化に向かっていく中で外部専門家がどのような役割を果たしてきたのかを理論ではなく実践に基づいて整理したものである。まず評価の目的を達成するためには、目的に合った仕組み、制度を作り上げることと、それを実際に庁内で動かしていくことの2つが求められる。その2つの観点から外部専門家の役割を整理すると、①評価の仕組みの設計と提示、②全体の仕掛づくり、③トップ幹部層に対する説得誘導、④関連部門の組織化、⑤職員に対する指導、⑥客観的な評価情報の提供、⑦住民と行政間の通訳と言った7点となる。そして外部専門家を活用していく中では、小規模自治体では外部支援を受けられない、手作りの仕組みは作ってくれない、外部支援とは外部評価と言った固定観念が自治体側によく見られる。それを払拭しないと正しい外部支援は実現しない。また今後の自治体評価における外部支援は、行政監査や職員能力開発との連動を図ることである。

キーワード

評価、仕組みと運用、改革改善、点検指導、会議運営

1. はじめに

筆者は1995年に三重県で行政評価制度の導入支援をして以来、13年間で250を超える自治体を対象に、外部専門家として評価全般に関ってきた。本稿は、日本評価学会の編集委員会からの依頼により、自治体の行政評価において外部専門家がどのような役割を果たしてきたのか、その役割発揮上での課題は何か、本来のあるべき姿はどうかという問題意識に対して、実際の経験に基づく整理を試みたものである。

よって、あくまで自らが関わってきた事例に基づ

く論拠が主体になっているため、参考文献立脚型の一般的な論文と比べると学問的根拠があるとは言えない。しかしながら本稿が少しでも今後の評価研究及び実践の一助になれば幸いである。

2. 評価の目的達成のための基本的考え方

外部専門家の役割を考察する前にまず、自治体が行政評価に取り組む際の基本的な考え方を述べる。行政評価制度を導入する目的は、各自治体でその表現は様々であるが、まちづくりの水準を高

め、行政活動の質を高め、効率的な遂行を行い、さらに住民などに対して説明責任を果たすということとされている。

こうした目的を達成するためには、その目的に合った最適な仕組みを制度として確立させることと、その仕組みを実際に庁内で動かしていくことの2点が基本となる。つまり、仕組み（制度）×運用＝目的達成、という方程式が成立する。

古川（2004）は、「評価の何が本物で何が偽物か」という論文の中で、自治体の評価活動が混迷に陥っている原因の一つとして、運用のまずさを指摘している。つまり、同じような評価の仕組みや方法を採用しても、実際に首長以下の職員が評価の本質を正しく理解して判断して行動しなければ何も変化は起きず、まさに評価作業疲れが蔓延してしまうことになってしまう。

上述の古川（2004）を受けて筆者は「本物の評価に至る視点」（2004）で、本物と偽物の評価についての考察を行い、それを見分けるチェックリストを添付し、本物の評価を実現するための評価の体系と視点を提唱した。このチェックリストは①評価制度の導入目的、②評価結果の活用、③計画策定や予算編成、組織の責任権限などの運用環境整備、④評価結果に基づく組織運営の流れについての判断基準から構成されているが、①については仕組みの範疇であるが、②以降の3点はいずれも運用の範疇での指摘をしている。

つまり評価制度の導入目的を確実に達成するためには、評価方法の選択や様式設計、指標設定など評価の仕組み構築だけでなく、むしろ、運用面での工夫、特に評価結果の活用やそのための関係各課の協力体制、評価結果に基づいて意思決定を行ない行動に結びつけるという動きがとりわけ、制度導入から定着化までの期間に必要となる。

3. 評価における外部専門家の役割

自治体が行政評価制度の導入の際に、外部専門家の支援や助言を依頼することは一般的となっている。田中（2006）によれば、都道府県と市・特別区の計285自治体の内、行政評価制度の設計において外部機関が支援や協力という形式で関った

ところが約67%を占めている。そして支援や協力を依頼された外部機関は民間シンクタンクやコンサルタントが75.8%を占め、一方で大学教授や公認会計士などの有識者が33.2%を占めている。

外部専門家に与えられた使命、1次的な役割は言うまでもなく、自治体の行政制度の導入目的を確実に達成していくことに他ならない。そしてその目的を達成するためには実務経験があって第三者的な立場で首長から職員、さらには議員や住民に対して助言や支援を行うことが求められる。これを仕組み面と運用面とで分けて役割として整理すると次の7点となる。

- (1) 評価の仕組みの設計と提示
- (2) 全体の仕掛けづくり
- (3) トップ幹部層に対する説得指導
- (4) 関連部門の組織化
- (5) 職員に対する指導
- (6) 客観的な評価情報の提供
- (7) 住民と行政間の通訳

この7つの役割を以降、順次解説していくが、1番目の「評価システムの設計と提示」が、「自治体の特性に合った仕組みを作りあげる」という仕組み面での役割に対して、他の6つの役割はいずれも、「どのように評価を進めていくか」という運用面での役割であることが特徴である。

この方程式が意味することは、要は、同じような仕組みを導入してもその運用によって目的を達成できる、あるいは逆に目的が達成されない、形骸化してしまう結果をもたらすということである。以降、各役割について解説する。

(1) 評価の仕組みの設計と提示

まず、評価制度の導入目的を確認し、それに合った適切な仕組みを設計し、それを首長以下関係者に提示することが第一の役割である。評価の対象が、政策体系における施策や事務事業であったとしても、それらの特性を理解し、評価者の評価能力に合わせて無理なく、かつ評価者が評価過程を通じて気づきと学びができるような評価の仕組みを設計することである。ただこうした仕組み自体はインターネットで容易に他自治体の事例を入

手できる時代なので、今ではあえて外部専門家に依頼しなくても自前で出来るのではないかと言う認識が自治体の担当者や首長に増えてきている。

ただしここで留意しなければならないのは、評価システムの設計イコール評価表の様式設計ではないことである。評価表の様式を設計する以前にまず、評価のロジックを組み立て、それが事務事業評価であれば事務事業の特性毎にロジックを確立させ、それを実践できるように単純化することである。単純化しなければその評価のロジックに従って実際の事務事業を評価してみることが出来にくくなる。

筆者は、1995年に三重県で全国初めての事務事業評価の導入に際して、その評価の仕組みを設計して提示した。それまで筆者は民間企業における組織業績制度の設計と運用について長く経験を有していたが、三重県からの評価制度導入の依頼に対しては、いきなり民間企業の組織業績制度を適用することをしなかった。むしろ行政の仕事である事務事業の特性を明らかにした。それは民間企業の組織や業務が比較的、目的や成果が明確であり、品質・納期・コストという定型化された指標で表されることに対して、行政の事務事業は目的が明確に設定されていない。事務事業の目的が最終目的と混在してしまっていること、また外部に対する補助金支給など間接的な性格を持つものが多いことなどを理解した。

そこで後年、多くの自治体が参考事例とした三重県の事務事業評価の仕組みが出来上がった。その評価ロジックは、まず目的と手段を分ける。目的を対象と意図で分けて明らかにしてその目的達成度を成果指標として定義し、さらに事務事業の上位の目的である政策体系の施策や基本事業の目的との峻別を行った。そしてコストと活動との関係を効率性、活動と成果との関係を有効性と定義して、評価のロジックをいきなりコストに注目した局所的なものではなく、政策体系を意識した目的妥当性評価、そして成果を意識した有効性評価、コストを意識した効率性評価、受益と負担の公平性評価と言う流れとして構築させた。

(2) 全体の仕掛けづくり

「全体の仕掛けづくり」とは、庁内の組織実態を踏まえて、実施計画策定や予算編成など既存の組織活動に対して、どう行政評価システムを連動させていくのか、管理職や一般職員をどの段階で啓蒙、理解推進、評価能力向上させていくのか等の工夫を言う。こうした工夫を具体的な組織の動きにしたものがマスタープランである。マスタープランは向こう3年から5年間を見越して各年度の到達目標、つまりどのような状態を作るのかを明らかにしてさらにその年度でどのようなことを組織的に行うのかを明らかにしたものである。

マスタープランを策定する際には、自治体組織の大きさや分業の仕方、総合計画の策定改訂年度や首長の任期など様々な要素を考慮しなければならない。よってこのマスタープラン策定には多くの支援経験が必要となる。この役割を果たす外部専門家は、言わば健康診断結果に基づいて処方箋や治療方針を示すドクター的な存在と言える。

三重県の場合、1995年に行政評価制度の導入支援を開始した頃、筆者が提示したものは先述した評価の仕組みだけではなかった。評価結果を確実に次の企画や改革改善、予算や定員管理などに反映させるための3ヵ年のマスタープランであった。行政評価の取り組みを契機にして、従来何をやっても中庸の県と言われた三重県の位置を、いかに先進県にしていくか、そのためには初年度からどのような到達目標を設定して、何をやるべきかなど組織体質改革を促すプラン作成と提示も重要な外部専門家の役割の一つである。さらにこのマスタープランはちょうど、山登りの際の案内人の様に、自治体組織が行政評価制度の導入目的達成のために確実に一步一步、動き始め、当面の到達目標を示してそれに向けて関係者を動機付けするものでなければいけない。

全体の仕掛けを考える際にまず必要なことは、政策体系の在り方を決める総合計画の策定期間がいつなのかとか、首長マニフェストと評価との連動をはかるために次の首長選挙はいつなのかといった点である。また合併をしたばかりで、事務事業の評価単位が不揃いだったりしていると、いきなり事務事業評価を行って予算に反映させる前

に、事務事業自体の単位の適正化の作業が必要となる。こうしたことも総合的に判断して、前進感のあるマスタープランを策定し導入目的達成に向けての仕掛けづくりを行う。

(3) トップ等の説得誘導

外部専門家の三番目の役割は、「トップ等の説得誘導」を行うことである。特に、首長が評価制度の導入にあまり積極的でない場合とか、議会における評価制度に対する理解が不十分である場合などは、トップや議会に対する説得誘導の役割は外部専門家では果たせない。それは職員の立場ではどうしても首長や議員を教えると言うことに抵抗を感じてしまうことと、受け手の側も決して聞く耳を持つ人だけではないからである。

また、首長が積極的に評価制度導入に取り組もうとしても幹部層、特に首長と幹部層による庁議が評価結果に基づく意思決定や方針を示すことが出来るように機能しないと、折角の評価結果を活用出来る運用環境が作れない。要は、従来多くの自治体組織で見られていた「報告連絡のための部長会議」、「意見が出てこない幹部会議」から脱却して、評価結果に基づいて議論を尽くして意思決定を行う会議へとレベルアップを行わなければならないのである。

そして、そのような庁議運営改革を行うためには単に、学識経験者による行政評価のトップ講演だけでは不十分である。長野県飯田市は、2002年より評価制度を導入したが、当初は理事者と部課長に対する学識経験者による講演会を開催ただけで、後は自前での取り組みであった。しばらくしてから行政評価の担当者は閉塞感を感じて筆者に相談にきた。それが契機となって筆者が市長、理事者、部長層に対して行政経営会議の運営に関する研修や助言を行うようになったことで、それまでの「評価のまねごと」からの脱皮が一気に進み、2005年度から2006年度に掛けて市民との協働による総合計画策定に反映され、2008年度からは政策体系の施策評価を市民交えて実践するというレベルにまで進化を遂げた。

このように単に「評価を知識として知る」ということと、庁議が経営会議として評価結果に基づ

いて意思決定を行うようになる、つまり「評価を実務として出来る」ということにはかなりの開きがある。そのための外部専門家の役割はさらに、単に講演や研修のみならず実際の庁議に参加し、議長代行で意思決定会議の運営をはかるか、あるいは第三者的な観点で評価ロジックに沿った助言を行い、庁議を活性化するかいずれかである。このいずれかの役割を演じるかは、首長、理事者と外部専門家間で事前に役割分担のすり合わせを行っている。筆者の場合、支援の当初は議長代行の役割を果たし、庁議運営レベルが向上し、理事者の中で意思決定型の会議運営が出来る人が出てきたら、今度は第三者的な立場で評価結果に対して意見を言う役割に変わることが多い。特に、首長の政治的配慮から開始された事務事業の評価結果を受けての廃止や抜本的改革への意思決定のための会議運営では、意識して納税者的な立場で意見を言うことが多い。また「この評価結果にも関わらず事務事業を来年度以降も存続するとなると納税者としての住民はどのように受け取るでしょうか？」等の質問を投げかけることもある。首長にとって耳の痛いことでも遠慮なく指摘するという崇高な姿勢が外部専門家に求められる。

(4) 関連部門の組織化

外部専門家の次の役割は関係部門の組織化をはかることである。とりわけ企画部門や財政部門、行革部門など行政評価に関わる部課が3つも4つもあって、縦割り意識が強い大規模自治体では、全体を束ねるオルガナイザー的な存在が必要となる。これも職員ではなかなか出来ない役割の一つである。関係部門に対して正論を言えないし、関係部門の担当者の方も正論を聞くことが出来ない。このことは組織の中で仕事をしている人から見れば当然のこととして受け止められている。

評価制度を導入していく際、特に、事務事業評価という単体の評価を超えて、政策体系の施策を評価単位として、その結果を次年度以降の経営方針設定に反映させていき、さらに予算編成や定員管理などの資源配分につなげていくには、一つの課だけの取り組みでは限界がある。こうした制度改革、経営の仕組み改革は、総合計画の在り方や

予算編成の在り方など既存の仕組みにも関わることである。

新潟県新発田市は、田中（2006）の調査結果において他の自治体が参考事例とした先進自治体としてリストアップされているが、2000年に事務事業評価に取り組んだ頃はある民間シンクタンクが支援をしていた。しかし、評価制度が実際に機能しなくて行政評価の担当者が問題意識を感じて筆者の元まで相談にきた。そこで筆者はそれまでの状況を把握し、新発田市が指向している行政評価の仕組みと運用は、局所的な事務事業見直しではないこと、最終的には政策体系に沿って予算編成の在り方自体を抜本的に変えることを確認した。そして次に提案したことは、総合計画担当、予算編成担当、組織定員管理担当、広報担当、人事評価担当など関係課の課長で構成される「新行政推進会議」であった。この会議を設置して、当初に市の経営問題を構造的に把握して関係課長間で問題を共有化したことで強力な推進母体となった。2003年には全国に先駆けて、部門単位ではなく政策体系単位での予算編成制度に切り替えた。

関連部門の組織化を果たしていくためには、外部専門家に求められることは計画や予算に関する知識よりもむしろ、評価と計画、予算を連動させることによって計画策定や予算編成の担当者にとっての問題が解決できるという道筋を示してあげることである。

（5）職員に対する指導

評価に対する職員の理解や能力を高めることは、行政評価に取り組む自治体にとって永遠の課題と考えられる。毎年の人事異動がそれをさらに強調させている。また自治体では特に、課長による部下の育成のための時間を確保することが難しく、職員の育成は自治研修所などの研修コース受講などによるOFFJT（職場を離れたトレーニング）が主体である。

そうした育成環境下であって、新たに事務事業評価や施策評価に取り組むとなると、評価表作成要領の配布と説明会だけでは不十分である。そこで評価制度の導入に際して外部専門家による評価研修とその後の評価結果の点検指導が最も効果的

な支援メニューとして広く多くの自治体に認知されるようになった。

評価結果の点検指導とは、評価担当者が行った評価結果に基づいて、評価責任者である課長や評価担当者に対して、その結果が正しく評価ロジックに沿っているかどうか等を確認し、必要な助言を行い、課としての意思決定に繋げる貴重な機会である。この点検指導が初めてわが国の自治体に採用されたのは三重県である。1996年4月に初めて導入された点検指導会方式は、1990年代から筆者が民間企業の経営指導において築き上げた新しい外部関与の方式である。それまでの外部の関わりは、シンクタンクや外資系コンサル会社にありがちな、ヒアリングを行なって、それを持ち帰り、立派な分厚いレポートを印刷製本してそれを納品し、せいぜい要約を幹部にプレゼンするというものであった。こうした外部関与の方式を筆者は「聞き取り報告書納品」方式と言っている。この方式をあらゆる場において否定するつもりはないが、とりわけ、評価制度を構築していく際には職員の正しい理解と、評価能力の向上が不可欠であり、そのためには点検指導会が最も有効と考えられる。

酒井（2005）は、ある自治体で点検指導会が開催された後に、それに参加した職員に対して行なったアンケート結果を発表してその有効性を証明している。非常に有効であったと回答した職員が36.5%、どちらかと言えば有効であったと回答した職員が52.7%を占めている。さらに点検指導会開催前と後で、事務事業評価結果とそれに基づく改革改善の項目が70%の増加となっている。なぜ、点検指導会が効果的と言えるかは一方的な講演や説明会だけだと受け身になってしまい、いやいやながら評価表をただ文字で埋めるということに陥りやすいからである。それに対して点検指導方式は、評価担当者が作成した評価結果をプロの視点で確認して的確な助言を行うものであり、双方向的であり、個別的でもあるので担当者の理解は格段に深まる。

点検指導会が活用されるのは、事務事業評価と施策評価の2つであるが、とりわけ数が多い事務事業評価の点検指導については、一つの事務事業に対して早くても90分、初回だと120分の時間が掛

かる。例えば外部専門家が経験12年の筆者であっても、それだけの時間はかかる。と言うのは受け手の側の担当者や課長が真に理解、納得するまでである程度の時間がかかるからである。しかもこの点検指導会では課長が参加することから課としての評価を決めるという意味合いもある。

酒井（2005）は、自らの経験から、点検指導を行う指導員としての外部専門家に求められる資質として、改革への信念、担当者を正しい評価ロジックに誘導するコーチング能力、議論を促すファシリテーション能力、担当者の悩みや問題意識を引き出すカウンセリング能力が必要としている。自治体の仕事の内容を知っているかどうかは二の次なのである。また評価についても単に知識として知っているだけではなく、評価ロジックに沿って現実の事務事業や施策をそこに乗せて、聞き出しながら判定していく評価実務能力こそが決め手となる。

田中（2006）は、行政評価に取り組んでいる285の自治体に対する実態調査の中で、職員の評価に関する理解や能力の改善が必要と答えた自治体が56.5%となっていることを明らかにした。その意味からも今後、自治体は評価に対する点検指導会の機会をもっと増やしていくべきと考える。そしてこの点検指導には酒井が指摘する通り、ベテランのプロの指導者が担当することが重要である。例えば最近の事例であるが山梨県南アルプス市は行政評価制度導入当初は学識経験者に支援を求めていたが、点検指導による職員の理解の重要性を考慮して、2007年度から筆者や酒井などプロのコンサルタントが点検指導に携わっている。

（6）客観的な評価情報の提供

外部専門家の役割の6番目は、評価担当者や意思決定者に対して客観的な評価情報の提供を行うことである。とりわけこの役割は、政策体系に基づく施策評価に関して重要性を持つ。施策評価では、まちづくりの水準に関する現状水準の把握を行うが、近隣他市との比較を行う際に例えば全国データを提供するなどの役割である。

ただこうした評価情報を提供する役割は、提供される側からすると安易に「こんな情報はない

か？」と情報過多となり、時には評価そのものを丸投げしてしまう風潮を引き起こしやすい。実際に筆者に相談に来る自治体の担当者の中にそのような姿勢になっている人が数人いる。

特に、評価に活用する指標の設定については安易に外部専門家が実態を知らない段階で評価者に対して「こんな指標がありますよ」などと言って情報を提供するのは危険である。指標に捉われすぎると指標に振り回されてしまい肝心の目的が見えてこなくなる。目的が明確でないまま指標を設定しても担当者は違和感を持つし、それに基づいて設定した目標値の妥当性も確保されない。

むしろより重要な役割は、メタ評価、つまり、行政側の評価がそのロジック通りに適切になされているかを、点検指導や評価会議運営などを通じて把握して、それらをとりまとめて評価者や評価担当部門、さらにはトップに対して報告し、適切な対応策を勧告するという役割になっていくであろう。こうしたメタ評価を行うためには外部専門家は、正しく評価を理解していなければならないし、評価の仕組みの問題なのか、あるいは運用の問題なのかを判断できる知見と問題解決能力が求められる。

（7）住民と行政間の通訳

外部専門家の最後の役割は、住民と行政との間に入って両者の意思疎通を良くし、特に住民に対して正しい評価の仕組みや運用に対する理解を促すことである。また同時に職員に対しても、住民の視点で評価することを具体的な場面で提示する。例えば職員の視点では「施設を整備する」と言う言葉を、住民の視点から「施設を利用する際にはこのような期待がある」と翻訳することがここで言う通訳としての役割である。

先述した飯田市では、2003年度から行政評価制度を本格導入したが、2005年度と2006年度の2ヵ年に渡り、市民と若手職員によるワーキングチームを編成して政策体系を再設定し、施策の現状と課題を明らかにして、施策の成果指標に基づく目標設定を行った。特に、施策の目標設定の際には、市民と若手職員で何度も議論を重ねて、単に行政に要望事項をとりまとめるのではなく、施策の成

果向上に際しての主体としての役割を明らかにした。例えば、災害対策の推進という施策では、一人ひとりの市民は災害に備えて3日間程度の飲料や食料を備蓄することや、地区では防災マップを作成して配布をすること、市役所の役割は公共施設の耐震化を図ることなど、多様な主体でまちづくりに取り組むという姿勢で計画が策定された。さらにその目標を達成するための各主体の果たすべき役割の発揮度を表す指標を「ムトス指標」と名づけた。こうした一連の協働作業に対して筆者を筆頭とする外部専門家が果たした役割は、市民と行政との間に入って相方が同じ方向を向いて、各々の役割を定義できるような翻訳家であった。

こうした翻訳家の役割を果たしていく上で、外部専門家に求められるのはオーケストラの指揮者のような組織化能力である。行政側の下請け的な存在でもなく、住民側の代弁者的な存在であってもいけない。時に住民に対しても、行政職員に対しても耳の痛いことを指摘しなければならないこともある。一方で、どのような進め方や演出をすれば彼らのやる気を高めることが出来るかというモチベーション策も求められる。

4. 外部専門家の活用に対する固定観念

田中（2006）の実態調査結果に基づくと、行政評価制度の導入に際して自治体に対して関わる外部専門家がここ数年増えてきている。そして先述した通り、外部専門家が果たすべき役割の重要性は今後さらに増大していくものと思われる。

しかしながらここに来て、いくつかの課題が明らかになっている。第一は、外部専門家に対する正しい理解がなされてないことである。自治体側にある種の固定観念が広がっており、まずはそうした思い込みを解消して正しい理解を促すことが必要と考える。

（1）小規模自治体では外部支援を求められないという観念

「行政評価の導入で外部の支援を求めるには膨大な費用がかかるので、小規模自治体にはその余

裕がない。行政評価は小規模自治体にはなじまない」と言った声がたまに聞かれる。この考えはまず、外部の専門家の役割は評価の仕組みを作ることであり、大きな自治体は比較的大きな予算を使えるので外部機関に委託は出来るが、小規模自治体はそれが出来ないということに立脚している。

ところが実は、小規模自治体は指導対象の課の数なり管理職数もそれだけ少ないから、少ない日数と費用で支援が可能となる。また庁内における関係課との調整の手間もかからない。一方で一般的に小規模自治体は首長が全体を見渡せるのでワンマン的な組織運営となり、管理職層の評価能力や説明能力が十分育ってないことが多い。よって管理職層の能力向上の上でも外部専門家の指導が必要ではないかと考える。

（2）外部の専門機関に依頼しても手作りの仕組みは作ってくれないと言う観念

二番目の固定観念は、「外部の専門機関は自らの自治体に合った手作りの評価の仕組みは作ってくれない」、さらに裏を返せば「評価の仕組みは自らの自治体に合った手作りのものにしなければならない」というものである。実際に、自治体の担当者の中には、こうした固定観念から、評価のロジックを無視した単に書きやすさだけを考慮した評価シートを自前で作ってしまうことが多い。事実、筆者に相談に来る自治体の大半は既に独自の評価シートを作成しており、そのことが足かせになって評価担当者が評価疲れをしており、改革改善への導きがない、現状追認型の評価となっていることが多い。

そもそもスポーツでも料理でもそこには基本の型がある。それを理解しないで局所的、近視眼的にシートを設計しても仕組みの導入目的は達成されない。例えば事務事業評価の場合、目的と手段を分ける、手段の大きさを表す指標を活動指標といい、対象における意図の達成度を表す指標を成果指標と定義する、事務事業評価は総合点で評価は出来ず、目的妥当性評価から入るロジック展開になる…などは自治体の規模や地域特性には関係なく基本的なことである。自治体担当者の実務経験に裏付けられない「手作り」によって決して歪

められてはいけない原理原則である。

むしろ自治体の規模や地域特性、首長や議会など政治特性などによって個別に考えなければならないのは仕組みではなく、どのように庁内を動かしていくかという運用の部分である。例えば行政評価制度の導入に際して、マネジメントサイクルとして関係する総合計画の策定時期や期間、首長の任期やマニフェストとの関係、議会の関わり方、財政事情や予算編成方針などどのように関連させていくのか、展開ステップや体制などは10の自治体に10通りのパターンがある。事実、筆者が外部専門家として支援している自治体では全く同じような進め方はない。

(3) 外部支援とは外部評価であるという観念

三番目の固定観念が、「自治体による自己評価は客観性が担保されないから、評価には市民団体やNPOが関わるべきだ。行政評価の外部支援にそうした団体が使える」と言うものである。言わば外部支援とは外部評価であるという観念である。

外部評価とは正確には自治体の職員以外の人が自治体の施策や事務事業を評価することを言うが、その前提となるのは自治体職員側がきちんとした評価をやっていることである。また外部評価といっても評価者がどのような立場で評価をするのかを予め明確に決めていないと、納税者的な立場の部外者は「もったいないから止めるべきだ」と言った厳しい評価を行う反面、受益者的な立場の部外者は「もっと充実すべき。つまり費用を増やすべき」となりがちである。これでは外部評価は行政側が恣意的に判断して委嘱した外部の人たちによる恣意的な判定結果に陥る恐れがある。

また仮に外部評価を行ったとしても、その結果を自治体側が真摯に受け止めて意思決定を行わなければならない、単なる評論に止まってしまう。よって自治体側の評価能力、問題解決能力、説明能力を向上させることが前提となる。その意味でも外部評価を行う前に、自治体側の自己評価の水準を高めて確保するための外部支援が必須であると言える。その外部支援は自治体内部に入り込んで第三者の立場から自治体職員に対して鋭い質問を与え、質疑応答の過程での気づきと理解を促すもの

でなければならない。

5. 外部専門家活用の留意点

自治体が評価制度を導入して定着化を進めて行く際に重要な点は、自治体側に主体性を持つことである。決して全てを丸投げするものではない。言い換えると、外部専門家を招聘する側の姿勢によって、外部専門家の活用範囲や方法も左右される。所詮、外部専門家は雇われ用心棒のような存在であることは否定できない。

新発田市も飯田市も、行政評価の導入担当者はこれまでの振り返りをした際に、外部専門家の必要性は認めたものの、「100の自治体に100通りの評価がある。主体性を持って庁内に働きかける姿勢が重要である」とか「支援を受ける側の主体性ある展開が必要」と異口同音に強調している。つまり、行政評価の担当者は常に庁内における改革の仕掛けを自らの頭で考え、外部専門家に相談をして、協議の中で具体的な対策を打ち立てるという仕事の進め方を採るべきと考える。つまり、外部専門家の限界もそこにあることになる。あくまで支援の立場であり、自らが首長や担当者にとって代わることは出来ない。

また記入要領などのマニュアル策定を請け負う外部機関もあるが、こうした支援形態の場合、マニュアル提示、評価作業の進行管理と言った形式的なものになりがちである。先述した通り、目的の設計なくして指標設定はありえない。いたずらに指標設定に埋没してしまい、自らの意図する状態が見えなくなるという恐れもある。評価に関する留意点は行政評価担当者が手作りで庁内にわかりやすく提示することが望ましいと考える。

6. これからの自治体評価における外部支援指導の在り方

これまでの実践的な取り組みを通じての気づきを主にとりまとめた。最後に、これからの自治体の行政評価において外部支援指導はどのようになっていくか、なっていくべきかについて述べる。

(1) 自治体評価の方向性

まず自治体における評価制度は、総合計画策定や予算編成、さらには人事評価に至るまでトータルの行政経営の仕組みとして位置づけられ、来るべき歳入減少時代に対応すべく歳出構造の徹底的な見直しの一環として進化していくと予想出来る。仕組み自体が個別の時代から統合の時代へと変るであろう。また目標設定自体が当たり前のようを受け止められ、まちづくり目標設定の際に近隣他市との比較も容易になっていくであろう。

また今後5年後を見越すと、団塊世代の部課長が一気に第一線から退くことになるので、40代あるいは30代の若手職員が急速なスピードで部課長に就任しなければならない状況下にある自治体も多い。そのような場合は、5年先を見越して幹部職員を早急に育成しておくことが重要課題となる。人材育成のスピードを今まで以上に加速していかなければならないのである。

(2) 今後の外部支援指導の在り方

そのような中で外部専門家による外部支援指導の必要性和期待は益々高まっていくと予想出来る。自治体側が外部支援指導を有効に行うためには、行政監査や職員能力開発との連動を図ることである。自治体の行政評価の仕組みや運用がその導入目的と比べて合目的であるのかと言った監査の視点を持って、その結果を首長に適宜、報告し適切な対応策、改善策を勧告することもより行政評価制度の目的達成のために有効な改善策である。

さらに職員の理解と評価能力、改革改善能力を高めるために各種能力開発プログラムとの整合性をはかり、職員のOTJ（職務を通じてのトレーニング）として外部専門家に関ることが望ましい。言わば、評価の実践と職員研修とを連動させることである。職員からしても実際の役所の仕事やまちづくり課題を考えながら、必要とされる能力の向上につながる事が出来る。

また、外部専門家自体の抜擢や育成も重要な課題となる。先述した通り、自治体評価に対する外

部専門家の役割は多岐に渡るし、さらに深い。その役割を確実に果たしていくためには評価の仕組みを理解しているだけではなく、より組織や人間に対する洞察力も求められる。学生や公務員、民間コンサルタントなど候補生を集めて、その適性を見抜き、必要な時間と経費を掛けて育成していく体制が求められる。

筆者は行政評価の他にプロフェッショナル人材育成をコンサルティングのテーマとしているが、単に知識修得レベルではなく、経験から学習する能力や、評価ロジックの応用展開能力、さらには安易に首長や職員に迎合することなく、客観的、公平公正な立場で指導支援をする崇高な倫理観など今後の調査研究の余地は大きいと思われる。さらなる調査研究の進展を期待して締めくくりたい。

参考文献

- 古川俊一（2004）「評価の何が本物で何が偽物か：その判断基準」、日本評価学会春季第1回全国大会報告論文集
- 星野芳昭（2004）「本物の評価に至る視点」、日本評価学会春季第1回全国大会報告論文集
- 田中啓（2006）「行政評価の実態に関する調査」調査報告書、静岡文化芸術大学特別研究論文
- 中村征之（1999）「三重が燃えている」、公人の友社
- 酒井隆（2005）「評価能力向上のための実践手法：点検指導会運営」、日本評価学会春季第2回全国大会報告論文集
- 古川俊一・星野芳昭・梅田次郎編（2003）「行政評価実践ゼミナール…行政経営改革のフロンティア」、ぎょうせい
- 星野芳昭（2001）「ガバメントガバナンスと行政評価システム」、公人の友社
- なお、新潟県新発田市の事例については、「新発田市における計画、予算、評価の連動」（2005年7月）を、長野県飯田市の事例については、「飯田市における基本計画策定過程と行政評価による進行管理」（2007年5月）と言う公式資料に基づいた。

（2008.1.23受理）

Outside Supporting Services for Local Governments in Introducing and Stabilizing the Performance Management System

Yoshiaki Hoshino

Renovation Producer, Director of JMAC Governance and Corporate Renovation Sector
hoshino@stones.com / yoshiaki_hoshino@jmac.co.jp

Abstract

The roles of outside professional services for Local government in introducing and stabilizing the Performance Management System (PMS) are not only showing evaluation methods but also the organizational renovation suggestions, auditing and guiding the results of performance measurement. This article emphasize on the right methods of PMS as well as the right orientation of organizational changes in utilizing the results of PMS based upon the long time pragmatic experiences.

Keywords

evaluation, system and its utilization,
renovation and improvement (kaizen) auditing and coaching, meeting managment

【研究ノート：依頼原稿】

非営利セクターと自治体の政策評価

岩渕 公二

参議院 政策担当秘書

lupin_xebec@nifty.com

要 約

本稿の目的は、岩手県を中心に評価制度の導入支援や運用支援、外部評価ならびに外部評価の中間支援など幅広い活動を展開してきた特定非営利活動法人政策21を事例として取り上げ、人員や予算、知識・技術などの資源が限られる地方の自治体の評価において政策評価NPOが果たしうる機能と課題を明らかにすることにある。そのため、最初に政策21の設立経過と組織的特長、活動内容を整理するとともに、政策21の活動フィールドとなった岩手県内における評価制度の導入状況等を整理している。次に政策21が実施してきた受託事業の内容から政策21が担ってきた役割を分析し、最後にそれぞれの関与局面において政策21が果たしてきた機能を明らかにしている。本研究により、地域に密着した政策評価NPOの柔軟な支援機能の有効性を確認するとともに、政策評価NPOとしての組織的・機能的課題を提起することができる。

キーワード

政策評価、NPO、外部評価、エンパワーメント、インボルブメント

1. はじめに

1990年代後半、全国自治体の評価制度導入の先駆けとなった三重県の事務事業評価導入から、10年が経過した。この間、全国の都道府県ならびに政令市、中核市、特例市の多くが評価制度を導入し、それ以外の市町村でも、評価への取り組みが積極的に続けられている。しかし、地方の小規模の自治体では、人材や予算、知識・技術などの資源の制約によって、評価制度の導入や運用に苦慮している現実もあり、いまだ外部からの支援に対する要請が続いている。

一方、すでに評価制度を導入した自治体では、

より能率的で実効性の高い評価制度をめざしてシステムの改善に取り組むとともに、客観性の向上をめざした外部評価の導入を志向し、様々な試みが続けている。そのとき評価者として台頭しているのが、非営利セクターの一つ、特定非営利活動法人（以下「NPO」）である。大学教員や民間研究機関の研究員、公認会計士や税理士といった士業者によって構成される委員会形式の評価とともに、評価の専門性を備えたNPOへの業務委託や協働による評価が取り入れられるようになった。

自治体の評価にNPOが関与してきた事例には、三重県の特設非営利活動法人コミュニティシンク

タンク「評価みえ」や静岡県の特設非営利活動法人コミュニティシンクタンクふじなどが挙げられる。前者は2000年に設立され、協働事業の評価に取り組んできたほか、NPO認証を受ける前年の1999年には「市民による事業評価システム99」を作成するなど、市民による評価の先導役を担ってきた。また、後者は2003年に設立され、同年度に富士市の業務委託によって補助金272件総額24億円を見直し、翌年度の予算に反映させた実績を持つ。

そうした中で、本稿で取り上げる特定非営利活動法人政策21（以下「政策21」）は、評価制度の導入・運用の支援、外部評価の実施、外部評価の中間支援と、前述したNPOよりも評価に関する活動の幅が広く、政策評価NPOとして注目できる。非営利セクターであるNPOが自治体の評価にどのようなかわり、どのような役割を果たしてきたのか。以下、政策21を事例として、自治体の評価における政策評価NPOの機能と課題について論じる。

2. 政策評価NPO

設立経過

政策21は、2000年12月に設立準備会、2001年1月に設立総会を開催し、同年2月に認証申請、同年4月13日に岩手県知事の認証を受け、NPOとして活動を開始した。設立時の社員（正会員）は、大学教員と社会人院生が中心で、本格的な政策評価NPOとして、また市民が立ち上げたNPOとして、国や自治体関係者、研究機関、報道機関等の関心を集め、地元マスコミをはじめ全国紙や雑誌でも取り上げられるなど、社会的反響を呼んだ¹⁾。

政策21設立の背景には、1998年4月に開学した岩手県立大学の存在があった。同大学では、県立大学として地域貢献を最大のテーマに掲げ、実学・実践の教育とともに、地域課題の解決に向けた学際的教育をめざして総合政策学部を設置し、当時としては全国に数少ない政策評価講座を開設した。同大学では、2000年4月、学部完成前に大学院を開学し、社会人教育に積極的に乗り出した。政策21の設立メンバーはその第1期生が中心で、

自治体職員、自治体議会議員、会社経営者、市民団体の代表等、多彩なメンバーが集まり、幅広い視点からの議論が可能な環境があった。

組織的特長

政策21の設立目的は、「政策評価の研究と実践、各種研究会やセミナーの開催を通じて政策評価制度の普及をはかり、政策の実務者だけでなく、市民の政策型思考と市民の代表である議会の政策立案能力を高めるとともに、積極的に政策提言を行うことによって、広く公益に寄与しようとする」ことである（政策21定款第3条）。その手段として、すべての政策分野における研究・実践活動を展開することとし（同定款第4条）、政策提言、事業受託、講師派遣やセミナー開催等の事業を行うこととした（同定款第5条）。そのため、政策21では、社員（正会員）に政策評価ならびに政策領域での専門性を重視して入会資格を限定し（同定款第6条）、具体的には「政策評価に関する研究領域で修士以上の学位」ならびにそれに準ずる実務経験との基準を設け（同理事会決議）、現在に至っている。

入会資格が限定的であったことから、無料で情報提供を受けられる情報会員以外の社員（正会員）は増えていない。表1は、政策21社員ならびに政策21の活動に主体的に参画できる有料会員の推移を表したものである。備考欄に示したように、現

表1 特定非営利活動法人政策21有料会員数の推移

年 度	有料会員		備 考
	種 別	数	
2001 年度 (2002 年 3 月 31 日現在)	正会員	13 名	大学教員 2 名含む。学位取得者 9 名（博士 1 名、修士 8 名）
	参加会員	7 名	国・地方自治体の職員、NPO 職員
	団体会員	3 団体	1 団体 2 社
2002 年度 (2003 年 3 月 31 日現在)	正会員	16 名	大学教員 2 名含む。学位取得者 10 名（博士 1 名、修士 9 名）
	参加会員	14 名	国・地方自治体の職員、NPO 職員
	団体会員	3 団体	1 団体 2 社
2003 年度 (2004 年 3 月 31 日現在)	正会員	16 名	大学教員 2 名含む。学位取得者 13 名（博士 1 名、修士 12 名）
	参加会員	16 名	国・地方自治体の職員、NPO 職員
	団体会員	3 団体	1 団体 2 社
2004 年度 (2005 年 3 月 31 日現在)	正会員	14 名	大学教員 2 名含む。学位取得者 14 名（博士 3 名、修士 11 名）
	参加会員	12 名	国・地方自治体の職員
	団体会員	2 団体	1 団体 1 社。※2004 年度をもって団体会員制度を廃止
2005 年度 (2006 年 3 月 31 日現在)	正会員	14 名	大学教員 2 名含む。学位取得者 14 名（博士 4 名、修士 10 名）
	参加会員	9 名	国・地方自治体の職員
2006 年度 (2007 年 3 月 31 日現在)	正会員	16 名	大学教員 2 名含む。学位取得者 15 名（博士 5 名、修士 10 名）
	参加会員	8 名	国・地方自治体の職員

（出所）特定非営利活動法人政策21の各年度事業報告書ならびにヒアリングをもとに筆者作成

在では社員（正会員）16名中、25年以上の経験を持つ民間コンサルタント1名を除く15名が修士号以上の学位を持ち、そのうち5名が博士号を持つ専門家集団となっている。

活動内容

政策21には、設立直後から自治体や公益法人、労働組合などからの事業依頼が相次いだ。その内容は、評価制度導入支援の業務委託や評価制度検討会へのアドバイザー派遣、職員研修やセミナーへの講師派遣等、幅広いものであった。図1は、政策21の活動内容を示したものである。政策21では、活動のキーワードを「交流」「支援」「協働」におき、行政機関、研究機関、民間団体や市民を対象に事業を展開してきた。

なかでも、自治体を対象とした事業の割合が高く、設立年度から現在までの6年半（2007年9月30日現在）の間に自治体から受託した事業数は、29件に及ぶ。その内容は、評価制度導入にかかる調査・研究、政策評価に関する基礎知識の講演、評価制度導入に伴う職員研修、すでに運用している評価制度の改善に関する助言、外部評価の実施、外部評価の支援等となっている。政策評価の専門性を組織的特長とする政策21の特色を活かした活動を展開してきたといえる。

次項では、政策21の活動と最も密接に影響し合

ったとみられる岩手県内の評価制度状況について論じる。

3. 岩手地区の政策評価

(1) 岩手県

政策評価システム

岩手県では、1997年に事務事業評価、1998年には公共事業評価を導入し、1999年の総合計画の策定を経て、2000年4月から県庁知事部局で政策評価システムを導入した。本庁での政策評価システムは2001年4月から本格実施に移され、2002年4月からは出先機関である各地方振興局でも評価制度の設計・試行が始まり、全庁的に政策評価に取り組んできた。

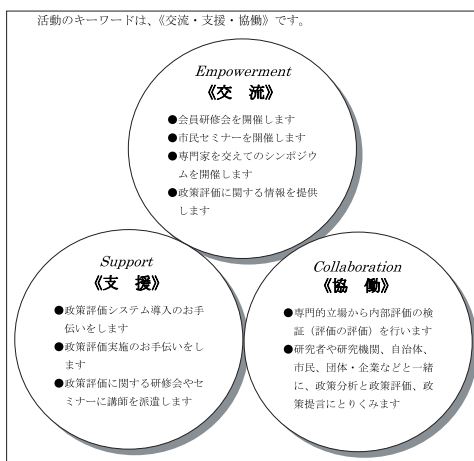
その主な目的は、総合計画の進行管理であり、総合計画に掲げた「施策」「分野」「主要な事業」「主要な指標」を対象として評価を行っている。評価はすべて内部評価で、「分野」「主要な事業」「主要な指標」は各部局、「施策」については評価専門部署が実施し、評価制度のチェック機能として政策評価委員会、ある一定規模以上の公共事業については公共事業評価委員会が設置されていた。これらはいずれも要綱・要領を根拠とした内部制度であったが、岩手県では2003年9月の県議会に条例案を提出し、2004年1月から条例に基づく政策評価システムを運用している。

政策評価制度の条例化は、宮城県など他の自治体でも取り組んでいるが、その主な目的は、制度に自治立法権に基づく根拠を与えて実効性をより高めることと、評価結果とその活用に正当性を持たせることにある。岩手県では、条例化を機に従来設置してきた政策評価委員会と公共事業評価委員会を統合して第三者機関を一本化し、研究者等を中心とした専門性の高い「岩手県政策評価委員会」を設置している。

外部評価

岩手県では、内部評価の充実とともに、外部評価にも積極的に取り組んでいる。千厩地方振興局が2002年度に振興局長が所管する二つの補助事業

図1 特定非営利活動法人政策21の活動



（出所）特定非営利活動法人政策21のリーフレットより

を対象に外部評価を導入し、政策21に依頼して実施した。二つの補助事業とは、地域活性化事業調整費と市町村総合補助金である。地域活性化事業調整費は地域の特性を生かした個性ある地域振興を図るために市町村等が地域課題の解決あるいは地域の活性化を推進する事業を行う場合に要する経費に対して補助する制度である。また、市町村総合補助金は地方分権時代に対応した市町村の自主的な地域づくりを支援するために市町村等が岩手県総合計画に掲げる「環境・ひと・情報」の三つの視点を踏まえた事業で、かつ市町村等の創意工夫による自主的な活用と認められるものを行う場合に要する経費に対して補助する制度である。そのうち評価対象となったのは2000年度と2001年度に実施した事業で、その中から実施年度および実施主体別に2事業ずつ抽出し、計14事業を評価した。

この事業で注目される点は、千厩地方振興局が内部評価の設計・試行を実施する前に外部評価を導入したことであった。評価は委託契約に基づき行われたが、評価対象とする事業の選定や評価手法は政策21が独自に行い、評価結果の取りまとめおよび公表資料の作成も政策21が単独で作成していることから、業務委託とはいえ第三者機関による外部評価として一定の機能を果たしたと考えられる。この点について千厩地方振興局では、内部評価を実施する前に外部評価を実施することにより評価の考え方や手法を素早く吸収することと、振興局が所管する町村と同じ土俵で評価を受けることで、いまだ評価制度が導入されていない町村から評価に対する積極的な姿勢を引き出そうとしたとコメントしている³。

千厩地方振興局の試みは、県全体へと波及した。前述した二つの補助事業を管轄する岩手県地域振興部地域振興課では、政策21が千厩地方振興局の外部評価で実施した手法を県全域へと拡大し、12地方振興局ならびに58市町村が実施する二つの補助事業に対して適用することにした。さらに内部評価の適正を判断し、評価の精度を向上・維持する手段として、地域活性化事業調整費事業（以下「調整費事業」）については外部評価を導入し、2003年度から3年間にわたり実施した。

2003年度の外部評価では、評価の対象となる

2002年度の調整費事業が地域活性化事業調整費事業評価実施要領（以下「実施要領」）改正以前の事業であったことから、実施要領改正後の評価手法に基づき外部評価することは困難であったため、抽出事業を対象とした評価・調査を実施することとし、盛岡地方振興局と宮古地方振興局管内の22事業を対象としたヒアリングによる評価と、その他10地方振興局の60事業を対象とした電話等による聴き取り調査を実施した。2004年度の調整費事業の外部評価では、評価の対象が2003年7月の実施要領の改正に適応した事業であったことから、2003年度に実施された全事業について各事業の事業企画調書と事業評価調査に基づく外部評価を実施することとし、改めて評価手法の開発を行い、17のチェック項目を設定して合計100点満点の点数表記で評価を行った。

マニフェスト評価

2003年4月の統一地方選挙では、改革派といわれる知事がマニフェストを掲げて選挙に臨み、当選を果たした。その先導役を担ったのが、いち早くマニフェストを発表した増田寛也岩手県知事であった。増田知事は当選早々、マニフェストの評価を明言した。マニフェストに基づく政策の進捗状況を毎年度評価し、その結果を県民に公表して自ら積極的に首長としての評価を受けるためであった。

マニフェストの評価において重要な点は、従来の行政手段として実施されてきた政策評価では果しえない政治的領域のアカウンタビリティに踏み込むということである。マニフェストの評価は行政機関による政策実施のプロセスだけでなく、政策の判断や政策の管理にかかわる知事自身の姿勢と行動が評価の対象となるためである。もちろん、その結果に対する判断は有権者が選挙によって下すことになるのであるが、政策の実行度や成果、政策実現に向けた首長の姿勢など有権者の判断に影響を及ぼす情報がマニフェストの評価には含まれることになる。

増田知事のマニフェスト評価は、内部評価と外部評価の2本立てで行われた。内部評価結果を自己評価として公表するとともに、県民の実感に基づく外部評価を実施し、双方の評価結果を次の企画立案、次年度の予算編成に活かすためであった。

自己評価に代わる内部評価結果については、マニフェストを掲げて戦った選挙の翌年度以降、毎年「評価レポート」として発表された。一方、外部評価は、地方分権研究会などによる全国的な政策評価システムの可能性の検討を経て、任期の中間年度となった2005年度から実施された。

県民参加型外部評価

マニフェスト評価に関連しての外部評価は、直接はマニフェストを岩手県の行政計画に落とし込んだ「誇れるいわて40の政策」を評価対象とし、業務委託で実施された。事業名は「県民参加型外部評価システム構築事業」で、公募形式で実施された。

2005年度の評価対象は、県庁が内部評価で試行を始めた「特定課題評価」と同一テーマで、「高齢者の介護予防施策」と「若年者の就業支援施策」であった。対象とするテーマの選定にあたっては、総合計画の策定後隔年で行われ、2005年度以降毎年度実施することになった県民意識調査の結果をもとに、県民の関心の高い分野から抽出したものであったが、いずれのテーマも増田知事のマニフェストに掲げられた緊急優先課題の一つ「雇用対策」と、重点施策の一つ「安心して暮らせる社会先進県」に位置づけられていた内容であったことから、マニフェストの達成状況の測定と課題抽出を意識した評価であったことは明らかであった。

2006年度の評価対象は、評価団体の任意とし、外部評価を希望する団体から企画提案を受けつけた。公募期間中に寄せられた企画提案は、8団体6テーマにのぼり、審査委員による審査・採択を経て実施団体と評価対象が決定された。2007年度も同様に、評価対象は評価団体の任意とし、外部評価を希望する団体からの企画提案に基づき、実施されている³。

なお、2005年度に評価実施主体となった政策21は、2006年度以降評価実施主体となる各NPO等の支援活動を役割とし、外部評価の中間支援の役割を担っている。

(2) 県内市町村

岩手県内の市町村における評価への取り組み

は、比較的人員・財源等の資源が整っている市部を中心に展開した。国や県が市町村合併を推進する中で、行財政改革を迫られていた町村部では、評価制度導入よりも合併事務を優先せざるを得ない状況があった。町村でいち早く評価制度導入に取り組んだ雫石町は、合併しない方針を決めていたため、評価制度導入による自立への道を選択したもので、2002年度から政策21の支援を受け、政策評価システムの構築に取り込んだ。

表2は、岩手県内市町村の評価制度導入の推移を表したものである。

岩手県が政策評価に取り組み始めた2000年度当初の普及率は、全体で10%以下であったが、岩手県が政策評価システムを本格化し、県の出先機関である各地方振興局で政策評価の試行を開始した2002年度以降急速に伸び始め、合併により市町村数が減っている実態はあるものの、2006年10月1日現在では半数近い市町村で評価に取り組んでいる。

一方、合併しない町村、合併したものの規模が小さく市制に移行できなかった町でも、評価制度導入の検討が続けられている。しかし、資源（職員数と財源）が限られ、組織的に評価へ取り組むことが難しい町村も多く、表2で示したように、町村における評価制度の普及率は2006年10月1日現在で31.8%と、市と比べ普及が遅れている現状がある。

近年政策21が支援した町に、旧湯田町と旧沢内村が合併した西和賀町があった。評価の専門的な

表2 岩手県内市町村の評価制度導入状況の推移

年 度	市町村の導入状況			備 考
	市町村数 (そのうち町村)	導入済み・試行中 (そのうち町村)	割合 (そのうち町村)	
2000年度 (2001年8月31日現在)	59	3	5.0%	
2001年度 (2001年7月31日現在)	59	5	8.4%	
2002年度 (2002年7月31日現在)	58	13	22.4%	※2001年11月15日大船渡市と三好町が合併。
2003年度 (2003年7月31日現在)	58	18	31.0%	
2004年度 (2004年7月31日現在)	58 (45)	21 (10)	36.2% (22.2%)	
2005年度 (2006年1月1日現在)	41 (27)	17 (7)	41.4% (25.9%)	※平成の大合併により町村減
2006年度 (2006年10月1日現在)	35 (22)	17 (7)	48.5% (31.8%)	※平成の大合併により町村減

(出所) 総務省発表の資料をもとに筆者作成

知見を取り入れるにしても、外部に委託する財源や、複数の職員を外部研修に派遣する予算が確保できず、評価制度導入を検討する十分な環境を整えられない自治体も見受けられる。自立した地域経営の重要なツールである評価システムの導入が最も必要な資源が限られる自治体ほど、評価システムの導入に必要な知識や技術の吸収に制約が出ている実態が垣間見えるのである。

4. 要請と役割の推移

(1) 要請の変化

活動初期

政策21への事業要請は、国の評価制度の導入時期や岩手県の政策評価システムの運用、県内市町村の評価制度導入に向けた動きと密接に結びついてきたといえる。表3は、全国的な動きと岩手県の取り組みをまとめたものである。前述の岩手県内市町村の評価制度導入状況（表2）と合わせてご覧いただきたい。

政策21が設立された2001年は、国においては前年の地方分権一括法の施行に続き、「行政機関の保有する情報の公開に関する法律」が施行され、

表3 全国・岩手県内の動きと特定非営利活動法人政策21の支援内容の比較

年度	全国的な動き	岩手県内の状況	政策21の支援内容
1997年 ～1999年	地方分権一括法の成立（1997） 情報公開法の成立（1999）	事務事業評価（1997.10） 公共事業評価（1998.9） 情報公開条例（1999.4）	
2000年 ～2002年	地方分権一括法の施行（2000.4） 情報公開法の施行（2001.4） 政策評価法の成立（2001.6） 政策評価法の施行（2002.4）	政策評価試行（2000.4） 政策評価本格導入（2001.4） 千歳地方振興局評価事業（2002.9） ※年内市町村の20%超が評価を実施	法人設立、制度導入支援・普及啓発が中心 外部評価のパイロット事業実施
2003年 ～2005年	ローカル・マニフェスト導入（2003.4） パーティー・マニフェスト導入（2003.11） 今後の行政改革の方針（2004.12） 集中改革プランの公表を要請（2005.3） 政策評価法の見直し（2005.12）	地域活性化事業調整費事業 外部評価（2003. 4） 県民参加型外部評価システム構築事業（2005.6） ※年内市町村の30%超が評価を実施	制度運用支援ならびに第三者評価の要請が増加 マニフェスト評価に連動した県民参加型評価システムの構築
2006年～		県民参加型外部評価システム構築事業（2006.6） 県民協働型評価（2007.5） ※年内市町村のほぼ半数が評価を実施	評価者から第三者評価の支援へ転換

（出所）筆者作成

「行政機関が行う政策の評価に関する法律」（以下「政策評価法」）が成立した年度であった。

一方、岩手県においては、前年知事部局で試行された政策評価システムが本格的に移動し（2001年）、出先機関である各地方振興局で政策評価の試行を始めた時期であった（2002年度）。県内市町村においても、国の政策評価制度の導入決定、ならびに県および出先機関による政策評価実施を踏まえ、評価思考の啓発と評価制度導入の機運が高まってきた時期であった。そのことは、表2で示した岩手県内市町村における評価制度の普及率を比べても明らかである。

このように、政策21の活動初期段階では、評価制度導入に関する要請が期待された時期であったことがわかる。

転機のきっかけ

政策21が設立して3年目を迎えた2003年、転機となる全国的なうねりが起こった。マニフェストの登場である。

2003年春の統一地方選挙では、全国11の知事選挙で15人の候補がマニフェストを掲げて戦い、6名が当選を果たした。マニフェストの要件は、目標・財源・期限・工程が明確にされていることであり、事後検証が可能であることを最大の特徴としている⁴。マニフェストをめぐる地方からのうねりは、その年国政へと波及し、秋の総選挙では各政党がマニフェストを掲げるに至り、その後政治の舞台で定着していった。その先鞭となり、原動力となった改革派知事の一人が、岩手県の増田知事であった。

また、2003年の統一地方選挙以降、地方から地方分権を推進する発言や行動が活発化した。2003年5月の地方分権研究会による「三位一体改革の実現に向けての緊急声明」の発表、同年8月の六県知事による補助金不要リストの公表、2004年10月には全国知事会による「三位一体改革に関する提言」ならびに地方分権研究会による「真の三位一体改革の推進に向けての緊急声明」の発表、同年11月には21世紀臨調知事・市町村長連合会議による「地方分権改革としての真の三位一体改革の実現に向けて」緊急アピールが発表されるなど、政治・行政の枠組みについての議論が活発になっ

た。すなわち、地方自らが国との関係を見直し、自立への道を主張し始めたのが、この時期であった。こうした地方自立への動きは、政策21の活動にも変化を及ぼしていった。

(2) 役割の変遷

導入支援から運用支援へ

表4は、政策21が設立年度から2006年度まで実施した受託事業等の一覧である。一覧では、不定期に実施されている自主事業については除外したが、公開で行われるセミナー等には、毎回総務省行政評価事務所の職員をはじめ、自治体職員や自治体議会議員、市民団体の関係者などが参加しており、地域における評価情報のリソースとして活用されてきたことがうかがえる。

表4 特定非営利活動法人政策21の事業とその機能

年 度	事業名	委託先	機 能
2001 年度	行政評価基本方針策定業務	岩手県水沢市	制度導入支援
	総合政策評価システム研修	岩手県一関市	制度導入支援
	管理職研修	山形県尾花沢市	制度運用支援
	職員研修	静岡県天竜市	制度運用支援
	平成 13 年度静岡総研セミナー	《財》静岡総合研究機構	普及・啓発
	政策評価講座	岩手県遠野市職員労働組合	普及・啓発
	NPO 自己評価シートを作成	《社》日本青年会議所 NPO 支援委員会	第三者評価支援
2002 年度	行政評価基本方針策定業務	岩手県水沢市	制度導入支援
	行政評価研修及び導入基本方針策定支援業務	岩手県雫石町	制度導入支援
	地域活性化事業調整費等評価事業	岩手県千厩地方振興局	第三者評価
	職員研修	静岡県天竜市	制度運用支援
2003 年度	政策評価監修	東管井地域活性化推進会議	普及・啓発
	政策評価導入支援業務	岩手県雫石町	制度導入支援
	地域づくり事業企画・立案・評価手法習得セミナー事業業務	岩手県北上地方振興局	制度運用支援
	政策評価システム改善業務	岩手県水沢市	制度導入支援
	地域活性化事業調整費外部評価等業務	岩手県北上地方振興局・宮古地方振興局	第三者評価
	職員研修	岩手県北上地方振興局	制度運用支援
	公開講座講座	岩手県八戸地方振興局	普及・啓発
2004 年度	公共事業評価研修	神奈川県横浜直市研修	制度運用支援
	政策評価推進支援業務	岩手県雫石町	制度運用支援
	地域活性化事業調整費外部評価等業務	岩手県地域振興部	第三者評価
	事務事業事後評価（試行）指導点検会業務	福島県原町市	制度運用支援
	環境省モデル事業評価手法策定業務	岩手県住田町	制度導入支援
	事務事業事後評価（試行）検証報告書並びに行政評価実施に係る提言書作成業務	福島県原町市	制度運用支援
	地域活性化事業調整費外部評価等業務	岩手県地域振興部	第三者評価
2005 年度	雫石町政策評価推進支援業務	岩手県雫石町	制度運用支援
	事務事業評価全体研修会講師派遣業務	福島県原町市	制度運用支援
	県民参加型外部評価システム構築業務	岩手県総合政策室	第三者評価
	事務事業事後評価（試行）事前研修会講師派遣業務	福島県原町市	制度運用支援
	事務事業事後評価（試行）事後研修会講師派遣業務	福島県原町市	制度運用支援
	環境と経済の好循環のまちモデル事業効果測定・評価業務	住田町森林・林業日本一の町づくり推進協議会（岩手県住田町）	制度運用支援
	行政評価研修会講師派遣業務	福島県南相馬市	制度導入支援
2006 年度	雫石町政策評価推進支援業務	岩手県雫石町	制度運用支援
	県民参加型外部評価システム構築事業に係るサポート業務	岩手県総合政策室	第三者評価支援
	環境と経済の好循環のまちモデル事業効果測定・評価業務	住田町森林・林業日本一の町づくり推進協議会（岩手県住田町）	制度運用支援
	行政評価職員研修講師派遣	岩手県奥州市	制度導入支援

（出所）特定非営利活動法人政策21の各年度事業報告書ならびにヒアリングをもとに筆者作成

2001年度から2003年度にかけては、評価制度導入支援が中心であった。継続して支援している雫石町の場合でも、2003年度までは「導入支援」で、2004年度以降は「推進支援」となっており、制度導入から運用へと支援目的が変化していることがわかる。また、2004年度から2カ年にわたり支援した福島県原町市（合併後、南相馬市）は、その内容が既存の評価制度の定着と推進のための研修・提言となっており、運用面での支援であった。自治体における評価制度導入の潮流、国の政策評価法施行を見据え、いち早く評価制度を導入した自治体からは、評価制度導入の効果を高めるための支援要請を受け、活動を展開してきたことがわかる。

支援から評価主体へ

政策21では、設立当初から外部評価にも取り組んできた。2002年度の千厩地方振興局における評価事業を皮切りに、2003年度からは岩手県地域振興部の調整費事業にかかる外部評価を担当してきた。千厩地方振興局の外部評価は振興局による評価制度構築に先駆けて外部評価を実施したことから、外部から内部への評価思考と評価スキルの注入を目的としたものであったが、地域振興部の外部評価は全事業の外部検証を実施したことから、客観性を重視した第三者評価の役割を担っていた。同じ外部評価の形式による事業でも、政策21が担ってきた役割は、事業実施時期と背景により変化してきたのである。

一方、2005年度から始まった県民参加型外部評価システム構築事業は、より多くの機能を担ってきた。そのキーワードは、「県民参加型」である。この点について岩手県の増田知事は、県民参加型外部評価システム構築事業にかかるシンポジウムの基調講演『自立した地域経営をめざして～外部評価で住民参加を』の中で、政策評価への住民参加は「究極の住民参加である」と語っており、政策の意思決定プロセスへの反映を強く意識していた意図がうかがえる。政策21では、2005年度に実施した外部評価の設計において、住民参加を強く意識しつつ、参加型評価の手法を取り入れ、適切な評価の実施に向けた評価手法の設計を行った。すなわち、県民参加型外部評価においては、「住

民参加」と「参加型評価」という二つの「参加」をキーワードにシステムの構築と評価が行われた⁵。

評価主体から中間支援へ

県民参加型外部評価システム構築事業は2006年度も継続され、2007年度は「県民協働型評価」と事業名を変更し実施されている。しかし、2006年度以降、評価主体の育成と評価への幅広い県民参加をめざし、政策21は評価主体から外部評価の実施主体への知見・スキルの助言・提供を目的とした支援業務に特化している。県民の外部評価の参加と、外部評価を通じた県民の県政への参画を促進するため、政策21の専門性を活かしつつ、県民が理解しにくく参加し難い専門性を迫及した評価に偏らない外部評価のシステム構築をめざしてのことであった。

その背景には、県民参加型外部評価を発案するきっかけとなったマニフェストの存在がある。マニフェストは首長と有権者の契約であり、マニフェスト実現のためには有権者、すなわち県民の関与は不可欠である。とりわけ自立した地域経営を実現するためには、地域の重要な政策アクターである県民の参画が欠かせない。専門性を確保することによって客観性を高め評価の正当性を担保しようとする第三者評価とは異なる、住民参加の仕組みづくりとしての外部評価の姿が浮き彫りになってくる。

5. 政策評価NPOの機能と課題

(1) 関与局面

これまで政策21が果たしてきた政策評価NPOとしての機能について、事業の推移と支援内容の変化から、関与の局面を分析する。

政策評価NPOを含め外部の機関が自治体の評価にかかわる局面は、「設計」「測定」「判定」「判断（選択・決定）」の4つの段階が考えられる。なお、本稿では、外部の機関が自治体の評価にかかわる局面として、情報を集め（collect）、分析し（analyze）、測る（measure）活動の一つの段階として捉え、有効性等評価の観点・基準に基づいた

判定を行うために必要な情報を収集・分析し、効果の出現状態を客観化する一連のプロセスを「測定」段階と定義し、論じている。

設計は、評価の目的や評価の時点、評価の観点、評価の基準といった評価制度の基本スキームの検討・決定と、評価に使用する評価手法および評価調書の検討・決定・作成を行う。測定は、判定の基礎となる情報の収集と整理・分析を行う。判定は、評価の観点として定めた項目、たとえば政策評価法の観点でいえば必要性・有効性・効率性・公平性・優先性ごとに、事前評価であればその予測の妥当性を確認し、中間評価および事後評価であればその実績を計画時と比較して客観化する作業である。判断（選択・決定）は、事業の取捨選択などを決める意思決定の段階である。評価の現場では、「判定」と「判断」が同時に行われる場面も少なくないが、本稿では「判定」の結果を踏まえ何らかの意思決定を行うプロセスを「判断」と定義し、区分している。

政策21の活動初期に多く要請された評価制度導入の支援は、設計段階の要請である。また、評価制度運用の支援は、測定と判定を主な段階とし、一部に設計の変更を伴う制度改善の機能も含まれる。一方、外部評価には、多くの場合第三者評価としての判定に最も期待が寄せられている。しかし、政策21が関与した岩手県における外部評価では、その機能が様ではない。千厩地方振興局の評価事業ならびに地域振興部の調整費事業の外部評価においては、判定の要請が最も強いものの、県民参加型外部評価システム構築事業においては、2005年度は設計・測定・判定のすべての段階で、2006年度以降は設計段階での要請が最も強いといえる。

(2) 機能

政策21が自治体の評価と関与してきた局面は、設計から判定まで幅広い。しかし、その機能は概ね二つに集約される。「エンパワーメント（empowerment）」と「インボルブメント（involvement）」である。

政策21の場合、受託事業における委託者とは良好な関係にあり、信頼も厚く、事業実施にあたっ

ては主体性を発揮できる環境があった。それは、委託契約書ならびに仕様書にも反映されており、委託者から受託者への関与を極力抑える内容となっているほか、報告書の著作権を共有とし、受託者側の公表を自由にするなど、受託者の自律性とイニシアチブを最大限に活かせる環境が整えられていた。すなわち、委託者と受託者は、業務委託契約ながらも対等な立場による協働関係を維持する工夫が施されていた。

協働関係において、受託者である政策21が委託者である自治体へ与える影響は、評価に関する知識や技術、情報などの移転であり、外部から内部へのエンパワーメント効果である。一方、県民参加型外部評価システム構築事業（2007年度は県民協働型評価）における政策21の果たした機能は、システム構築の過程では関係者の参画を促すインボルブメントであり、外部評価への参加者に対して政策評価に関する知見とスキルならびに評価対象分野に関する県政情報の理解を深めてもらうエンパワーメント機能、受託者に対して評価結果の報告ならびに政策提言を通じて評価対象分野に関する課題や県民ニーズ、課題解決のアイディアなど提供するエンパワーメント機能である。

(3) 課題

政策評価NPOの先駆けとして、岩手県における市町村の評価制度導入や外部評価に貢献してきた政策21ではあるが、いくつかの課題も提起している。

一つは、評価の専門性と「わかりやすさ」の問題である。自治体の内部評価は、科学的根拠に基づく客観性を追求する志向が強いため、評価作業は精緻化し、県民にわかりにくい評価になっている。この傾向は、岩手県においても同様で、外部有識者で構成される評価委員会でも、専門的な議論を尽くことで客観性を高めようとしている。その結果、政策評価の理論だけでなく、土木や環境など個別の専門的な議論に終始し、委員会の結論との因果関係が県民に理解しにくい内容となっている。すなわち、具体的な議論の中身より、専門家が議論したという形式に対する信頼から県民の理解を得ようとしている。

しかし、県民が期待するNPOなどによる外部評価は、県民にわかりやすい評価、県民の生活実感に基づく評価である。専門性を高めているからこそ委託者である自治体の信頼を得られ、評価結果に対する期待感とともに、評価結果の活用可能性が高まることになるが、評価において専門性を追求し過ぎると、評価過程も評価結果もわかりにくいものになり、透明性を疑われる結果になりかねない。

もう一つは、政策評価NPOの組織の維持と社員構成の問題である⁶。政策21の場合、前述したように社員（正会員）に入会資格を設け専門性を維持してきたが、その結果社員（正会員）の数が伸び悩み、事業実施に制約が生じて多くの要請に応えられない状況もうまれた。

また、政策評価においては、多様な政策分野が対象となることから、政策分野それぞれの専門性の確保や、市民の多様な価値観を反映するための工夫が必要となる。この点について、政策21では、県民参加型外部評価システム構築事業で青年会議所や各地のNPOなど県内の民間団体に参加を呼びかけ外部評価プラットフォームを結成し、県民の多様な価値観を反映させる仕組みづくりを行ったほか、大学教員や民間研究機関の研究員等で構成するアドバイザーボードを設置し、評価対象分野の専門性を確保する取り組みを行っている。

入会資格を拡大すれば、社員が増え、運営の幅は広がるものの、専門性の維持が課題となる。一方、専門性を優先して入会資格を絞れば、社員が限定され、活動が制約される。このジレンマをどのように克服し、組織を維持していくかが、運営上の課題として挙げられる。

6. むすび

岩手県における政策21の活動の背景には、当時政策評価講座を開設していた岩手県立大学の存在があり、改革派知事として地域の自立を志向していた増田知事の存在があったことは否定できない。その意味では、すべての自治体に汎用性のある事例とはいえないかもしれない。

しかし、いまや地域経営の重要なアクターであ

るNPOや市民団体が、人・モノ・金・情報の乏しい地方において果たすべき役割は大きい。評価を地域経営のマネジメントツールとして位置づけるならば、外部評価には内部評価とは異なる機能を求め、役割分担することも有効であろう。行政機関による内部評価や行政活動の確認・検証・チェックの機能を果たす「行政監視型外部評価」から、政策推進の機能を発揮できる「政策推進型外部評価」の活用へ。第三者評価を含め、外部資源を活用した評価の仕組みを地域経営に活かすためには、その担い手となる人材と組織の育成が重要になる。

政策21の活動ならびに果たしてきた役割を俯瞰したとき、新たな制度導入のみならず、既存制度の改善や見直しといった柔軟な支援機能が注目される。カスタマイズされたシステムではなく、セミ・オーダーメイドやオーダーメイドのシステム開発、既存システムを円滑に動かすアプリケーション装置としての支援機能は、地域に密着した、小規模ながらも専門性を備えた集団で、かつ営利目的でない組織だからこそ為し得た技かもしれない。

その一方で、組織の維持や運営に制約が生まれ、活動の幅が広がらなかったことも事実である。全国各地に同様の政策評価NPOが育ち、互いに連携してネットワークを組み、それぞれが地域に根ざした活動を展開することが望ましい。

その意味で、2006年度以降の外部評価で政策21が果たしている評価の中間支援機能は、自治体の評価にかかわる市民・NPO等の育成に貢献するものと期待を寄せている。

注記

- 1 2001年2月6日付の朝日新聞は、社会面で認証申請の事実を報道した。そのほか、認証申請した同年2月5日から同年4月13日の認証までの間に、共同通信をはじめとする各紙・各局で取り上げられたことは、当時の社会的反響を物語っている。
- 2 千厩地方振興局の評価事業に関する同局職員のコメントは、2003年日本総合研究所「創発自治体Web」のホームページに掲載された（現在は閉鎖）。
- 3 外部評価事業ならびにその結果、評価結果の反映状

況については、岩手県ホームページ <http://www.pref.iwate.jp/> 参照。

- 4 日本総合研究所調査部経済・社会政策研究センターが2003年10月24日に発表した「自民・民主両党のマニフェストの比較 - SMART基準からの評価分析 -」の中で、イギリスにおける政策目標の5つの評価基準をそれぞれ「法律制度・改正を明示しているもの」「改革のための新組織創設を謳っているもの」「数値目標が明示されているもの」「政策のコスト・財源が明示されているもの」「政策の実施機関や期限が明示されているもの」と定義し、評価を行っている（日本総合研究所調査部経済・社会政策研究センター 2003）。
- 5 国民生活審議会総合企画部会がまとめた「我が国経済社会におけるNPOの役割と展望」（2001年4月9日）では、「国民ニーズの多様化、従来のコミュニティの役割の縮小などから、NPOの社会的意義が高まっている。市民が個人としての時間を使い、行政や企業がこたえられない社会の多様なニーズに対応することが重要」と位置づけている。多様化する社会では、公的サービスの担い手として市民や市民セクターに期待が寄せられているが、評価への住民参加はそうした消極的事情ではなく、主権者として、社会の構成員として、活動や運営の結果に責任を持ち、自ら地域社会を運営するという積極的な姿勢に基づく住民参加と位置づけられる（山谷・岩渕 2001）。
- 6 NPOには、制度、組織および活動形態に起因する特性が認められる。これらが評価プロセスや関与レベルと連動し、評価にかかわるNPOには多面的で複雑なアカウンタビリティが存在し、時にはそれらが交錯する事態も見受けられる（岩渕 2004）。

参考文献

- 岩渕公二（2004）「評価にかかわるNPOのアカウンタビリティ—地方自治体における市民参加の新たな潮流—」日本公共政策学会2004年度年報『公共政策研究』4、有斐閣
- 岩渕公二（2007）『外部評価の機能とその展開 - 行政監視と政策推進 -』、第一法規
- 窪田好男（2005）『日本型政策評価としての事務事業評価』、日本評論社
- 南島和久（2005）「行政活動を評価する」、佐藤竺監修

『市民のための地方自治入門 行政主導型から住民参加型へ』、実務教育出版
 日本総合研究所調査部経済・社会政策研究センター
 (2003)「自民・民主両党のマニフェストの比較－SMART基準からの評価分析－」
 山谷清志・岩渕公二(2001)「中央政府と自治体の政策評価」、財団法人三重社会経済研究センター編『地域政策－あすの三重』No.2、公人の友社
 山谷清志(2000)「評価の多様性と市民－参加型評価の可能性－」、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
 山谷清志(2001)「政策評価とスクラップ－参加型評価とNPO」、月刊地方自治『職員研修』34(6)
 山谷清志(2002)「政策評価とNPO－もう一つの実験」、日本公共政策学会2002年度年報『公共政策研究』、2、有斐閣
 山谷清志(2006)『政策評価の実践とその課題－アカウンタビリティのジレンマ』、萌書房

(2008.1.30受理)

Collaboration in Evaluation Systems of Non-Profit Sector and Local Governments

Koji Iwabuchi

House of Councilors, The National Diet of Japan; Legislative Staff
 lupin_xebec@nifty.com

Abstract

The goal of this research is to clarify functions and issues that can be fulfilled by policy evaluation NPOs in the evaluation of local governments that have limited resources, such as human resources, budgets, knowledge, technology etcetra. This was done by examining as a case study the Non-profit Organization POLICY 21, which has developed a wide range of activities such as evaluation system implementation and operation support, external evaluation and external evaluation intermediate support, primarily in Iwate Prefecture. To this end, in addition to describing the initial process of establishment, structural characteristics, and activity details of POLICY 21, the implementation status of evaluation systems in Iwate, the POLICY 21 activity field, is also described. The role taken on by POLICY 21 is analyzed based on the details of commissioned projects it has implemented, and the functions performed by POLICY 21 at various stages of participation are clarified. This research confirmed the effectiveness of flexible support functions in policy evaluation NPOs that have close regional ties, in addition to raising organizational/functional issues of policy evaluation NPOs.

Keywords

Evaluation, NPO, External evaluation, Empowerment, Involvement

【研究論文】

市場化テストのPDCAサイクルに見る評価手法の考察 ーマニフェスト評価体系からのアプローチー

田中 弥生

大学評価・学位授与機構

ytanaka@niad.ac.jp

工藤 泰志

(特) 言論NPO

y-npo@mta.biglobe.ne.jp

要 約

本論の目的はハローワークの市場化テスト事業という具体的な題材を通じて、現行の政策評価の問題を考察し、その解決の方向性を示すことである。市場化テスト制度のもとでは事前、事後の評価の仕組みが整い、PDCAサイクルに基づいて運営しやすい環境が整えられている。日本ではハローワークの市場化モデル事業が実施されたが、政治的な影響を受けてPDCAが機能しなかった。だが、この問題を現行の政策評価では説明することはできない。そこで、政治レベルの政策決定プロセスを評価対象に含めるマニフェスト評価アプローチを用いて分析することで問題の所在を明らかにした。そして、マニフェスト評価と政策評価の体系の差異を分析することで、現行の政策評価の課題と解決の方向性を考察した。

キーワード

政策評価、マニフェスト評価、市場化テスト、ハローワーク、PDCAサイクル

1. はじめに

政策評価は「各府省が、自らその政策の効果を把握・分析し、評価を行うことにより、次の企画立案や実施に役立てるものであり」、さらに、「政策評価の機能はPlan（企画立案）、Do（実施）、Check（評価）、Action（企画立案への反映）という政策マネジメント・サイクルの働きとして考えることができる」と説明される（総務省 2007）。そして、その目的は行政業務の効率化とともにアカウンタビリティ、すなわち国民に対する行政の説明責任を果たすことだとしている。

しかしながら、この政策評価の体系は府省の上位にある政治レベルでの政策決定プロセスを評価

の対象から外してきたため、いくつかの矛盾を生じされることになっている。

田中は、日本の構造調整借款の開始時期の政策背景に日米貿易摩擦に取り組む日本政府の葛藤と資金還流措置があったことを指摘し、府省の上位にある政治的決定によって策定された構造調整政策像を既存の政策評価やODA評価アプローチでは説明できないことを指摘した（田中 2005）。

山谷と田辺は「アカウンタビリティの対象である問責者の存在が見えず、答える責任を果たす方法の指定がなかったために」アカウンタビリティを確保するメカニズムがうまく機能していないことを指摘している（山谷 2006、田辺 2005）。それは政策の決定過程が政策評価の対象から除外され

ているために生じている。つまり、政策はそれを提起する政党の公約が投票という民主主義的な手続きで合意されるところからその実行のプロセスが始まる。そのプロセスの全体像を評価の対象にできず、実質的に行政業務の評価の範囲にとどめていたことが政策評価自体の制約にもなり、府省が行った政策評価結果に対して直接的な判断を下すことはできない状況を招いている。さらに言えば、府省が行う政策評価制度のもと、政策の必要性や有効性などが評価されおり、同制度は一定の役割を果たしている。だが、この制度の枠組みの中では、政策の妥当性について国民が直接判断を下すことはできない。政策評価のアカウンタビリティの対象を国民としながらも、国民は傍観者として位置づけられているので、山谷が指摘するようにアカウンタビリティの対象は実質的に不在なのである。

このような状況下では、PDCAサイクルをバックボーンとして政策評価を機能させることについて疑問が抱かれる。政策実行途中で、府省の上位にある政治的な影響を受けても、現行政策評価によってこの点を明らかにできない。つまり、府省の上位による政治的な決定によって、政策評価結果が無視され、その結果、府省がマネジメントする政策のPDCAサイクルが崩れることがありえる。そして、仮にそのような事態が積み重ねられれば、明確な説明がなくとも府省の政策の上位にある目的の変更が可能になり、制度や施策の運営に影響を与え、結局のところ国民に対して何ら説明責任を果たしていないことになる。

本論は以下2つの点を目的とする。第1に、政策評価を実施したにもかかわらず、PDCAサイクルが機能しなかった市場化テスト制度によるハローワーク事業に着目し、その状況を政治レベルの政策決定プロセスを評価対象に含めるマニフェスト評価手法によって明らかにすることである。第2は、現行の政策評価とマニフェスト評価体系の構造の差異について比較分析し、政策評価のあり方を考察することである。

2. 市場化テスト：政策評価の対象

まず、分析対象として市場化テスト制度を選択した理由を述べておきたい。市場化テスト制度の下では、事前と事後の評価方法が一貫しており、PDCAサイクルをマネジメントする環境が整っているからである。だが、市場化テストモデル事業の対象となったハローワーク事業については、PDCAサイクルがうまく機能していなかった。この2つの理由から市場化テスト制度を分析対象とした。以下、市場化テスト制度と実施状況について説明する。

2.1 日本における市場化テストの試み

(1) 市場化テストとは何か

市場化テストとは、ある公共サービスについて官と民が対等な立場で競争入札に参加し、価格・質の面で最も優れたものが、そのサービス提供を担う仕組みで、その目的は公共サービスの経費の削減と質の向上にある（内閣府公共サービス改革推進室 2006）。本制度は、構造的な不況に陥った英国、米国などの政府が導入したNPM（新公共経営）の流れの中で生み出された行政改革手法のひとつである（本間 2005）。

この制度は2つの側面をもっている。第1の側面は競争入札によって当該サービスを担う行政府機関はそのサービスの運営のあり方を見直すことによって公共サービスの効率性と効果を向上させることである。第2の側面は従来、官が担ってきた公共サービスを民間開放することによって新たなビジネス・チャンスを生み出すというものである。

(2) 日本における市場化テスト制度導入の経緯

日本において市場化テスト制度導入の議論が本格化したのは、小泉政権下、規制改革・民間開放推進会議（宮内義彦議長）での議論からである。2004年3月同会議より、市場化テスト導入するための調査・研究を2004年度中に実施する提案が出され、閣議決定された（規制改革・民間開放推進

3カ年計画)。同年4月には「経済財政運営と構造改革に関する基本方針(骨太の方針)」において、市場化テストを2004年度中に制度設計、2005年度に試行的導入を行うことが閣議決定される。2005年度より、3分野8業種において市場化テストモデル事業が開始され、2005年度、2006年度の2年間にわたりモデル事業が実施されてきた。モデル事業は市場化テスト本格導入を前提とし、その評価結果をもって本格導入のあり方やガイドラインを見直すことが想定されている。その意味でモデル事業は市場化テスト制度運営のためのPDCAサイクルを機能させるための手段であったといえよう¹⁾。

2006年5月参議院で「競争の導入による公共サービスの改革に関する法律案(閣法第34号)」(通称 市場化テスト法)が可決、成立された。

2006年9月安倍政権が発足するが、同政権下においても積極的に市場化テスト制度推進策が進められている。2007年6月に経済財政諮問会議から提出された「基本方針2007」において、市場化テストの抜本的拡大に向けて施策が打ち出されている。

(3) 市場化テスト運営体制と評価の仕組み

2006年度より、市場化テスト運営にあたり、「官民競争入札等監理委員会」が法律²⁾に基づき設置された。同監理委員会は、市場化テスト実施課程において中立性、公正性を確保するために設置された。そして、①内閣総理大臣による基本方針の策定、②国の行政機関等の長等による実施要項の策定、③国の行政機関等の長等による官民競争入札の落札者の決定の過程において、監理委員会の議を経るものとされている。さらに審議の結果の開示、国や民間業者、地方公共団体などから意見を聴き運営方針に反映することなどもその役割だとされる(公共サービス基本方針 平成21年12月)。したがって、官民競争入札等監理委員会は市場化テストに関わる利害関係者の間に位置し、同制度を運営し、その実施プロセスで得られた情報をもとに改善・発展させてゆくためのフォーカル・ポイントとしての役割を果たすことが期待されているとすることができるだろう。

また、市場化テスト制度には、明確な評価シス

テムがビルトインされている。すなわち、入札段階において、府省は対象となった事業の業績については、KPI(Key Performance Indicator)と呼ばれる指標を選定し、指標にもとづいて従来の業務を評価し、契約における要求水準を定める。また、業務にかかるコストを貨幣換算し予定価格を決める(官民競争入札等監理委員会 2006)。したがって、事前段階において当該事業の評価が業績とコスト面から定量的に行われることになる。

また、契約後の進捗状況や事業終了後の成果については、先のKPIにもとづき測定され、またコストもチェックされる。また、市場化テスト導入の前と実施後の業績やコストを比較するだけでなく、類似条件下で同一事業を営む官の業績とコストとを比較して、市場化テスト導入の効果を客観的に測定しようとしている。

したがって、事前評価と同一のKPIに基づき事後評価が行われるので、事前から事後まで一貫した評価が行われ、また市場化テスト導入効果を定量的に測定することもビルトインされていることからPDCAサイクルがより機能しやすい環境を整えているといえることができる。

また、市場化テスト制度全体の見直しについては内閣総理大臣の責任であることが法律に規定されており、制度全体の評価と改善も法律上担保されている(競争の導入による公共サービスの改革に関する法律 第七条7)。本制度全体の運営に携わる官民競争入札等監理委員会も報告または資料の提出を求めることができるとされている(第四十五条)。その意味で、評価とPDCAサイクルを担保する環境を制度上、整えようとしていることがわかる。

(4) 市場化テスト実施状況

2005年度、2006年度に実施されたモデル事業はハローワーク関連4事業、社会保険庁関連3事業、行刑施設関連1事業の3分野8事業である。2005年度には23事業が、2006年度には33事業が民間委託されている。2005年度(平成17年度)と2006年度(平成18年度)を比較してみると、2006年度において、ひとつの企業が複数地域の事業を受託する傾向があることがわかる。

平成19年度に入り、市場化テスト制度は本格導入に入るが、総務省科学技術研究調査、厚生労働省ハローワーク関連事業、社会保険庁関連事業、(独)雇用・能力開発機構のアビリティ・ガーデンおよび「私のしごと館」運営事業、(独)情報処理推進機構の情報処理技術者試験事業が対象となり公示された。また、2007年5月9日には柳澤前厚生労働大臣が官民競争入札等監理委員会に対して都内2箇所の職業紹介窓口業務を市場化テストの対象にする旨の説明を行っている。

(5) 事後評価結果

2005年度モデル事業実施後、事後評価書が所轄庁のホームページで公開されている。2007年2月、社会保険庁より3事業について受託者の契約遂行状況（要求水準達成状況）と同規模地域における官側の業績とコストを比較した評価報告書が公開された。

ハローワーク関連の評価報告書は柳澤前厚生労働大臣が都内2箇所での施行の意思を表明した5月9日時点では公開されておらず、2007年6月1日に公開されたが（朝日新聞 2007.6.2）、公開された報告書の日付は3月20日であった。また、行刑施

表1 市場化テスト事後評価結果（ハローワーク関連）

ハローワーク関連事業

業務	地域	業績（業務）	コスト
「キャリア交流プラザ事業」	北海道：キャリアバンク 埼玉：ブライトキャリア 東京：ジェイマムチェンジコンサルティング 愛知：ブライトキャリア 京都：オムロンパーソネル	就職率：民間実施地域46.4-61.9%。 55%以上の目標を達成したのは5地域中2地域で国の方が上回っている。 定着率も国が上回っている。 利用者の満足度：アンケート回収率において民間が68.2%、国実施地域が85.2%と大きな差がある。 国における満足度がやや高くなっている。	費用：委託費に対する実際の費用の割合を見ると、74.1-119.5%で委託費と実費の大きな乖離はない。 定着就職1件あたりの経緯は国実施地域がやや安くなっている。
「若年向けキャリア交流プラザ」	大阪：学生援護会（株）	就職率：44.9%で、自ら設定した目標を達成できなかった。 定着状況80.9%。定着状況を勘定した就職率は36.4%である。（中高年地域と比較し低めである。） 満足度75%で目標80%に達してない。	経費：就職1件あたりの経費は9.6万円、定着就職1件あたりは11.8万円。委託費は10.9万円
「求人開拓事業」	北海道札幌地区 キャリアバンク 秋田中央地区 廣済堂 福岡北九州地域 ブライトキャリア	企画書に掲げた目標を達成した民間事業者はなかった。求人開拓推進委員ひとりあたりの実績を見ると、開拓求人件数、求人数が比較対象の国以上の水準になったものの、もっとも意味をもつ開拓総数に係る指標についてはいづれにおいても民間実施地域が国実施地域の水準に及んでいない。	国に比べ民間において割高となっており、特にもっとも重要な指標である充足数一人あたりのコストでは、民間は国の3倍程度を要するという結果となっている。

（出所）厚生労働省 市場化テスト評価委員 市場化テスト モデル事業実績評価（17年度）より筆者作成

設に関する評価報告書の公開を確認することはできなかった。

評価結果の内容は次のとおりである。

社会保険庁関連業務については、業種や地域によって要求水準達成度や業績に差異があるが、全体に要求水準を達したところが多く委託費ベースの経費削減効果はあった。経費削減効果は主として人件費削減によるものが大きく、戸別訪問による収納業務から電話督促に切り替えることによって可能になった。但し、督促し易い対象が優先され困難な対象や山間過疎地域などコストの嵩む地域が外されるなどの問題、あるいは受注者の赤字で補填されており、今後の継続性が憂慮されるなどの問題も指摘されている。

ハローワークの評価結果は官民の差異がより顕著で総じて芳しくなかった。キャリア交流プラザ事業、若年キャリア交流事業、求人開拓事業のいずれも、自社が設定した目標値に達せず、また官との比較でも業績は劣っていた。またコスト面でも国に比較して民間のほうが割高になっており、特に求人開拓事業において民間は国の3倍程度の経費を費やしていた。一部の違反行為についても指摘されている。

また、評価報告書には記されていないが、メディアおよび官民競争入札等監理委員会⁴から、入札不調の問題も指摘されており、2007年3月現在において、求人開拓事業において全国3地域で落札者が決まらず結局、官が担うことになった⁵（朝日新聞 2007.3.29）。

2.2 問われる市場化テストハローワーク事業のPDCAサイクル

日本における市場化テスト導入から、事後評価結果まで一連の流れをみてきたわけだが、ここで大きな疑問にぶつかることになる。すなわち、モデル事業を柱としたPDCAサイクルが機能していなかったのではないかという疑問である。事後評価結果から、ハローワーク関連のモデル事業の結果が良くなかったことは明らかである。それにもかかわらず、柳澤前厚生労働大臣が、2007年5月9日にハローワークの本体事業への市場化テストの本格導入として都内2箇所の無料職業紹介所にお

いて官と民の窓口を併設するかたちで民間を参入させることを発表している。

また、発表された時点では、モデル事業の評価結果は公開されていなかった。しかも、官民競争入札等監理委員会も評価報告書を所持していなかった⁶。同監理委員会の議事録には評価結果について議論した記述はなく、少なくとも公開された情報の中からは、評価結果について議論した事実を見出すことはできなかった。

つまり、厚生労働省はモデル事業の評価を実施しながらも、次期立案において反映されず、モデル事業結果の検討を経ずして本格導入が決定された。その結果PDCAサイクルが機能しなかったのである。この事実は、PDCAサイクルを主軸として、その主たる目的を国民へのアカウンタビリティとする現行の政策評価のあり方に対して重要な問題を示唆している。なぜならば、府省の上位にある政治レベルの判断が影響を及ぼしているのならば、それを含めた評価体系を持たない限り、対象の評価は完結しないからである。

そこで、この問題を明らかにするべく、政治レベルの政策および政策決定プロセスを評価対象とするマニフェスト評価のアプローチを採用してみたい。

3. マニフェスト評価視点からみた市場化テスト

以下、分析対象であるハローワークの市場化テストモデル事業について民営化の経緯に遡りながら概観し、次に分析アプローチとしてのマニフェスト評価手法を説明する。その上で、ハローワークの市場化テストについて分析を行う。

3.1 分析対象：ハローワークへの市場化テスト導入の背景と経緯

(1) 民営化・民間委託議論の経緯

ハローワークの民営化を巡る議論は1996年11月に発足した行政改革会議に遡る。国の行う事務・事業の民営化・民間委託の推進が議論され、その

対象として検討すべき業務のひとつとしてハローワークが行う職業紹介事業、雇用保険が取り上げられた。これに対して、労働省は民営化案に対して強い反対の意向を示した。

すなわち、職業紹介業務について一部の民間参入の事実を認めながらも、それは高額な料金のもと特定の専門職、技術職に限定されるもので、無料の職業紹介はセーフティー・ネットとして必要であり官が担うべきだと主張したのである。そして、民営化は、ILO88号条約に違反すること⁷、職業紹介事業と雇用保険制度を一体運営していることを理由に不可能であると労働省は説明している（行政改革会議 労働省説明資料 1997.5.7）。

2001年4月に発足した総合規制改革会議において「民間参入の拡大による官製市場の見直し」が議論され、民間参入の拡大の検討対象の12重点検討事項のひとつとして、ハローワークの行う「職業紹介」が取り上げられる。厚生労働省は、先に挙げたILO88号条約違反、職業紹介事業と雇用保険制度の一体化運営を理由に反対意見を堅持している。

(2) 市場化テスト導入議論

2004年4月に発足した規制改革・民間開放推進会議において、再び、ハローワークは議論の対象となり、ここで市場化テスト導入案が提示される。同会議は2004年以来、厚生労働省に対して、ハローワーク関連業務の包括的民間委託を行うことを提案している。これに対して、これまで民営化・民間委託に強く反対していた厚生労働省は、初めて方針を変え、同年12月24日に「第一次答申」を提示し、就職コンサルティング事業を実施する「キャリア交流プラザ事業」に関する市場化テストのモデル事業実施に同意した。また、2005年12月21日に提示した「第2次答申案」では「人材銀行」について市場化テストを実施し、そのために必要な職安法の特例を「公共サービス改革法」に盛り込むことに厚生労働省は同意している（規制改革・民間開放推進会議 2005.12.21）。

2005年より実施されたハローワーク関連業務⁸の市場化テストモデル事業が実施されたが、厚生労働省は包括的民間委託に対しては難色を示し続

けていた。だが、2007年5月9日、柳澤前厚生労働大臣は経済財政諮問会議（座長・安倍晋三内閣総理大臣）において、都内2ヶ所の無料職業紹介窓口（職業紹介、職業相談）を市場化テストにかけるとを表明した。

3.2 マニフェスト評価：分析アプローチ

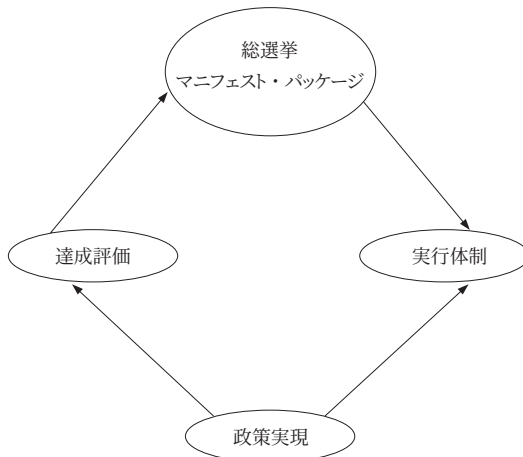
3.2.1 マニフェスト評価とは何か

マニフェストとは、有権者と政権担当者との契約を意味している。政党が政権獲得した暁には実行すべき具体的な政策パッケージを選挙前に有権者に提示し、政権獲得後はそれを実施し、政権期間中にどれだけ達成したのかを自己評価し、次の選挙でその判断を有権者（市場）に問う形でフィードバックされる。そこでは第三者評価も行われ、それらを含めて有権者の判断材料となる。ここでは政策を決めるのは政治と有権者の約束という関係であり、投票という民主主義の仕組みであることを前提としている。

こうしたマニフェストが日本の政治に導入されたのは、2003年の衆議院選選挙からである。英国の選挙で交付されるマニフェストを参考に従来の政権公約が抽象的なスローガンやwish listであるとの批判から、それを改め政策とその実現に向けた目標、さらには目標に対応する指標、財源、目標達成時期が具体的に明記されるのがマニフェストである（曾根 2004）。

マニフェストはそれを政策パッケージとして提示し、実施し、政策目標の達成度について評価を行い、次の選挙前にそれを有権者にフィードバックするというサイクルを成して完成することになる。図1は、マニフェスト・サイクルを示したものである。マニフェスト・サイクルには政策達成評価が組み込まれているが、政策に対する責任の所在を明らかにし、有権者へ具体的かつ明確にフィードバックをするという意味で評価が重要な役割を果たしている。つまり、現政権が掲げたマニフェストの実績評価結果と次期選挙に向けて提示したマニフェストの事前評価結果を判断材料にして有権者は投票を行う。政党はこうした有権者の行為を意識するので、有権者との約束であるマニフェストの実行と充実に尽力するようになる。こ

図1 マニフェスト・サイクル



(出所) 神吉信之の選挙コラムHPを参考に作成

うした有権者と政党の一連の行為が循環することによって、政治レベルでのPDCAサイクル（マニフェスト・サイクル）が機能するという考え方である。

マニフェストに掲げられた政策の達成度を評価するためには、その対象をマニフェストから、政治レベルでの政策決定過程、政策実施過程および実施結果までを含めて分析する必要がある。つまりマニフェスト評価は従来の政策評価が評価の対象から外した府省の上位にある政治レベルの政策決定過程を分析対象の範囲に含み、政策の導入は、選挙を通じて決定する有権者へのアカウンタビリティとして行われるものとなる。総務省が掲げた「府省が実施する政策評価」は、マニフェスト評価においては、その実行過程と結果の評価の一部として包含されることになる。

3.2.2 マニフェスト評価基準

(1) 言論NPOマニフェスト評価アプローチ採用の根拠

日本では、国政レベル、地方レベルの双方で複数の団体がマニフェスト評価に着手している。国政レベルでは、21世紀民間臨調が主催するマニフ

ェスト評価大会が著名であるが、2007年7月の参議院選挙を前に開かれた第3回大会では9団体が独自の評価視点に基づき評価結果を発表した。この団体の中で評価基準と評価体系を提示して評価結果を発表したのは言論NPOだけであった⁹。同NPOは、自民党、公明党、民主党が公開した21分野¹⁰のマニフェストの評価を行い発表している。また、小泉政権第1期から全てのマニフェスト評価を実施し公開したという実績を有する。

そこで、市場化テストのPDCAサイクル問題を分析するにあたり、言論NPOの評価基準を採用することにした。

(2) 評価基準と評価対象

ここではまず言論NPOのマニフェストの評価基準を説明する。表2は評価基準を要約したものである。①政策課題の妥当性、②実行プロセス、③実績、④アカウンタビリティの大きく4つの評価項目が設定され、①～③の項目はさらに形式面と実質面に分かれる。形式面と実質面に区分したのは、政策を提示し、審議会が設置され形式的な要件が整っていても、その内容が課題に忠実に向かい合っていない場合があるため、実質要件を設定することにした。政策課題の妥当性においては、政党が提案する公約の中身自体を対象にしていることがわかる。また、実行プロセスにおいては審議会を含むその運営体制、予算などの投入やそのプロセスが、公約で掲げた目的を実現する方向で適切に向かい合っているのかをみる。そして実績においては、アウトプットとアウトカム¹¹の2つのレベルを想定している。アウトプットレベルでは、当該政策が基本方針（骨太方針）など閣議決定文書に明記されていること確認し、アウトカムレベルでは所轄庁の実施結果と①で掲げたビジョンや政策目標との整合性を観てゆく（言論NPO 2007）。公約の評価と実行プロセス、さらに結果は一連の公約実行のプロセスを構成するもので、このサイクルに評価の体系が対応する形となっている。

表2 言論NPOのマニフェスト評価基準

要件項目	小要件項	要件視点	調査対象：ハローワーク市場化
政策課題の妥当性	形式要件	ビジョン・ミッション 課題設定、目標設定の明示 工程・時期の明示	マニフェスト 政権構想 所信表明演説、施政方針演説 進路と戦略 基本方針2006
	実質要件	ビジョンから政策までの政策体系 課題設定の妥当性 目標、施策体系の妥当性など	
実行プロセス	形式要件	課題を実現する体制 工程監理	経済財政諮問会議議事 ハローワーク民営化関連議論 官民競争入札等管理委員会議事
	実質要件	課題への取り組み状況など アウトカムを出すための指導力など	
実績	形式要件	アウトプットの判断で、閣議決定文書 に明記。政策実現のための道筋の明示	基本方針2007、市場化テスト法文書 官民競争入札等管理委員会議事 厚生労働省の市場化テスト導入方針 モデル事業事後評価報告書 受注企業からのヒアリング
	実質要件	アウトカムの判断 政策実行による成果の発現 成果とビジョン、目標との整合性	
アカウントビリティ		結果やプロセスを有権者がわかる形で 公開しているか。	上記プロセスにおける説明内容 公開状況

(出所) 言論NPOのマニフェスト評価報告書より筆者作成

3.3 マニフェスト評価によるハローワーク市場化テストの分析

以下、先のマニフェスト実績評価基準を基本にハローワーク市場化テストのPDCA問題に焦点を当てて分析を試みる。

(1) 政策課題の抽出のあり方

この評価でまず、確定しなくてはならないのは政策で実現すべき目標である。この目標は政策で掲げるビジョンと整合性を持ち、かつ測定可能な形で明示される必要がある。

そのため、安倍政権が自身の政権構想や政権獲得後の国民への説明となる所信表明、さらには施政方針演説などで、市場化テストをどのように位置づけたのか、どのように目標設定したのかが問われる¹¹⁾。安倍政権は政権構想「美しい国」の具体的政策の「自由と規律でオープンな経済社会」の第1番目に掲げた「官と民との新たなパートナーシップの確立」の中で、「小さく効率的な政府

の推進。民間活力フル活用」のために市場化テストの活用をすることを示唆し、所信表明（2006年9月）においても同様に記している。

さらに「進路と戦略」（2007年1月25日 閣議決定）において、民間の活動領域の拡大、効率的な行政の推進のための民間活力の最大活用という2つの政策目的において、市場化テストの活用を謳っている。ただし、政権構想に記されていた「官と民との新たなパートナーシップの確立」は1月25日閣議決定の「進路と戦略」以降、姿を消している¹²⁾。

また、小泉前政権は規制緩和・民間開放の手段として市場化テストを推進しているのに対し（規制緩和・民間開放会議 2004）、安倍政権は行政の効率化、民間開放の2つの目的を掲げ、その達成手段として市場化テストを据えている。だが、この変更とその理由に関する説明を公開情報から見出すことはできない。

こうした目的や目標は形式的に示されるだけでなく、実質的に記述される必要がある。そこでは目標の妥当性を判断するために、モデル事業実

施によって把握された課題を反映して進められたものかが問われる。だが、先の2つの目的を打ち出した2007年1月、およびハローワークへの本格導入の検討を首相が指示した2007年4月時点において、モデル事業の評価結果は公開されておらず、また官民競争入札等監理委員会にも提出されていなかった。

安倍政権がマニフェスト・サイクルに基づいて自ら打ち出した政策を運営するのであれば、モデル事業から得られた課題を抽出し、そこから行政の効率化、規制緩和・民間開放という2つの目的を実現するための達成目標やその期限を明示し、市場化テストの結果、先の2つの目的がどこまで達成されたのか評価を行う必要がある。だが、モデル事業をこうした政策目標実現のために活用するための議論はなく、新たに設定された2つの政策目標とモデル事業実施結果との論理的な結びつきを見出すことはできなかった。つまり、モデル事業の結果から導き出された課題にもとづいて市場化テスト事業を立案しているとは言い難い。

(2) 実行プロセス

実行プロセスにおいては、経済財政諮問会議および厚生労働省が、ハローワークへの市場化テスト本格導入を決定するにあたり、どのようなプロセスを経たかが問われる。また、各府省が行う市場化テストの調整役である官民競争入札等監理委員会がこの議論にどのように関与したのかが問われる。

ここではまず、経済財政諮問会議および厚生労働省の決定プロセスに着目する。

①ハローワーク市場化テスト導入を巡る経済財政諮問会議の議論

2006年11月、経済財政諮問会議は「ハローワークの市場化テスト導入について」という提案書の中で、市場化テスト導入の必要性を述べた上で、ILO88号条約については、最近のILOは民間職業紹介事業の評価や官民協力の必要性を認識しており、民間の補完が否定されるものでないと説明している。さらに、市場化テスト導入のための新提案として、現行の主要な官のハローワークを維持

したままで、その他の運営を民間に包括的に委託し、民間開放したハローワークを官が監督する仕組みを整えることで、官のネットワークは維持される、と述べている（経済財政諮問会議 2006.11.30）。

これを受けて、大田弘子担当大臣（経済財政諮問会議担当）は、ILO条約問題については、専門家による集中的な検討を考えると発言した（2006.11.30）。大田大臣の発言を受けて、大臣の私的諮問懇談会（内閣府公共サービス改革推進室）として「ハローワークとILO条約に関する懇談会」が開催された（2006.12.19）。同懇談会はILO条約違反について賛否両論あることを説明しながらも、諮問会議の提案はILO条約違反ではないと説明している。（ハローワークとILO条約に関する懇談会報告書 2007.4.6）（第7回経済財政諮問会議 2007.4.6）。

2007年4月第7回経済財政諮問会議においてハローワーク業務の市場化テスト導入について、諮問会議の官民双方の議員と柳澤前厚生労働大臣（諮問会議臨時委員）との間で激しい議論が展開されている。民間議員は、包括的民間委託はILO条約違反ではなく、無料職業紹介のネットワーク機能を確保したまま、無料の職業紹介業務に市場化テストを導入すべきであり、「骨太方針2007」までにその結論を出すべきと強く主張している。

これに対して当時の柳澤前厚生労働大臣は、1996年より述べてきた従来の理由に加え、ワーキングプアなど弱者のための福祉行政機関とハローワークとの連繋の重要性を理由に反対を表明している。これに対して民間議員はILOを組織防衛の道具にしてはいけない、「できない理屈の大天才」だとして柳澤前厚生労働大臣の対応を批判している。その後も同様の議論が繰り返されたが、最後に首相が「ハローワークへの市場化テストの導入をどのように行なうのか厚生労働大臣に考えて欲しい」とのコメントで締めくくっている（経済財政諮問会議議事主旨 2007.4.6）。

そして、2007年5月9日、柳澤前厚生労働大臣は諮問会議において、ハローワーク市場化テスト案として、都内2ヶ所の無料職業紹介窓口業務（職業紹介、職業相談）を民間委託部門と併設することを表明した。

ここで明らかなのは、経済財政諮問会議が2006年11月に「ハローワークへの市場化テスト導入について」の提案を皮切りに、厚生労働省が導入に抵抗を示したにもかかわらず、経済諮問会議が強くそれを推し進めようとしていることで、これは議事からも伺える。だが、この過程でモデル事業評価結果に関する議論は一切なく、事後評価報告書の公開は市場化テスト導入が決定された後の6月2日に公開されている。したがって、市場化テスト本格導入にあたり、モデル事業とその評価結果は議論の外に置かれてしまっていたということがいえよう。また、事後評価委員会には経済財政諮問会議の民間委員が委員として参加しているが、そうであれば評価結果が芳しくないことは事前に承知していたはずである。さらに、厚生労働省の上位にある経済財政諮問会議の委員が同省の評価委員になるという仕組みのあり方にも疑問が抱かれる。

②官民競争入札等監理委員会の役割

市場化テスト法に基づけば、官民競争入札等監理委員会は市場化テスト制度運営の調整役を担うことになっている。ハローワークへの本格導入決定プロセスにおいて同監理委員会はどのような役割を果たしていたのか。官民競争入札等監理委員会はハローワークへの市場化導入を決定的にした第7回の諮問会議には参加しておらず¹³、諮問会議の議長である安倍首相が、直接所轄庁大臣に指示を出している。法律上は、官民競争入札等監理委員会が市場化テスト制度運営を担うとされながらも、ハローワークへの導入に関する決定プロセスには直接的に関与していないと思われる。

また、同監理委員会は事後評価報告書を所轄庁から徴収する権限は制度上担保されているもののそれを行ってはいなかった¹⁴。柳澤前厚生労働大臣がハローワークへの本格導入を決定した時点において、同監理委員会は事後評価報告書を所持していなかった。さらに評価報告書が公開された後、官民競争入札等監理委員会の議事録には入札不調に関する議論、民間受託業者からの指摘を踏まえ、業務環境を改善するなどの議論があったことを確認することはできたが（官民競争入札等監理委員会 2007.7.26）、評価結果について直接議論

した内容を議事録から見出すことはできなかった。

つまり、この時点において、官民競争入札等監理委員会はモデル事業から得られた教訓や課題をもとに制度を運営している状況にはなかったと思われる。また、経済財政諮問会議の議事において、モデル事業やその評価結果に関する議論は確認されていない。そのような状況下でハローワークに関する主導権は経済財政諮問会議に掌握されたかたちになっていたと考えられる。

(3) 実績

市場化テストについては、既に前政権時代に法律が制定され、モデル事業が行われていることから、実績としてモデル事業実施結果に着目することにした。

2.1 (5) において、2005年度、2006年度に厚生労働省によって実施された事後評価結果を示したが、そのポイントは次のとおりである。

- ・就職率、定着率、顧客満足度のいずれにおいても受注者自ら設定した目標に達しなかった。
- ・官との比較においても受注事業者の業績は劣っていた。
- ・コスト面でも官との比較において受注事業者が割高になっていた。
- ・受注事業者の中には困難な対象を後回しにするなどの行為がみられた。

しかしながら、これらの評価結果は発注側である厚生労働省が行ったものである。そこで、受注側である企業にヒアリングを行ったが¹⁵、次のような点が明らかになった。第1に発注価格の問題が挙げられる。間接経費を削除し、人件費を非常勤職員程度の価格設定に落としても予定価格を下回ることができなかった。受託した事業についてもランニングコストの面で不足が生じており、これらの証言からは発注側が見積もった価格が過度に安価である可能性もあり、発注価格の妥当性を検証する必要があるという点である。第2は仕様書の問題で、仕様書に記されていた業務内容や備品管理の見積もりが不十分であったという点である。第3に受注企業からのフィードバックや意見

を徴収する仕組みがなかったことである。また、事後評価結果が出る前に次期入札がかけられていた。

発注側と受注側の立場の違いによって発言内容に食い違いが認められるものの、双方ともにモデル事業の実施結果が芳しくなかったことを認めている。

だが、次期入札は受注の現場で起こった問題や事後評価結果を反映せずに行われている。つまり、事業実施の現場レベルにおいてもPDCAサイクルは十分に機能していなかったといえるだろう。

(4) 発見事項

マニフェスト評価アプローチによる分析結果から次のようなことがいえるだろう。

第1に本格導入決定に直接的に影響を与えたのは経済財政諮問会議である。ハローワークの民営化・民活をめぐる議論は10年来の政治的な争点であったが、その解決を急ぐように経済財政諮問会議が本格導入決定に圧力をかけていた。しかしながら、同諮問会議の場では、モデル事業やその結果について議論されることはなかった。PDCAサイクルに基づいて市場化テスト制度を運営するための手段であったはずのモデル事業は、経済財政諮問会議の場で議論の外に置かれてしまい、PDCAサイクルは機能していない。その直接の原因は経済財政諮問会議がこうしたPDCAのプロセスを軽視し¹⁶、ハローワークへの市場化テストの本格導入を強硬に急いだ点にあるといえる。さらに経済財政諮問会議の民間委員が厚生労働省の評価委員に加わっており、報告書の内容を知る立場にしながら、評価結果が芳しくなかった点をまったく議論の遡上に載せずに、本格導入を急がせていた。

これらのことは、制度上はPDCAサイクルを担保する環境を整えていても、府省の上位にある政治的な決定によって、PDCAサイクルが壊される可能性があることを示唆している。

第2に、PDCA問題は厚生労働省だけに留まっていなかった点である。先の実行プロセスと実績の分析から明らかになったように、ハローワーク

事業の実施現場に事業の実施現場、市場化テスト制度全体を運営する立場にある官民競争入札等監理委員会では事後評価結果に基づく検討が行われていない。また、この時点では内閣総理大臣（内閣府）による評価も行われていない。モデル事業実施から2年ほどしか経過していないために十分な時間を確保していないこともあるだろうが、本格導入が決まった平成19年5月の時点では、事業実施現場、厚生労働省、官民競争入札等監理委員会、内閣府、そして経済財政諮問会議のどのレベルにおいても、評価結果と次期政策（計画）が連動しておらず、PDCAサイクルが機能していなかったといえる。

4. 政策評価とマニフェスト評価体系の差異に関する考察

本論は市場化テスト制度のPDCAサイクルの問題の所在についてマニフェスト評価体系を借用して分析を試みようとした。そこで本セクションでは、政策評価制度の基本構造を省みたうえで、その構造から市場化テストのPDCAサイクル問題を捉える。そして政策評価制度とマニフェスト評価アプローチの差異を分析した上で、問題解決の方向にマニフェスト評価アプローチが何を示唆しているのか考察したい。

(1) 政策評価制度の基本構造

①政策評価法と主要ステイクホルダーズ間の関係

政策評価は平成13年（2001年）に「行政機関が行う評価に関する法律」（以下政策評価法と呼ぶ）に規定された。第1章第1条および第2章、そして第4章第2条に「政策評価の客観的かつ厳密」という言葉が記されており、客観的で厳密な評価を推進し、実施を維持することが重んじられた制度であることがわかる。

また、平成14年（2002年）4月より3年間の施行を経て、平成17年（2005年）に、本制度の見直しが行われ同年12月に「政策評価に関する基本方針の改定」が閣議決定された。改定のポイントは大きく4点あり、・内閣の重要施策に関する評価の

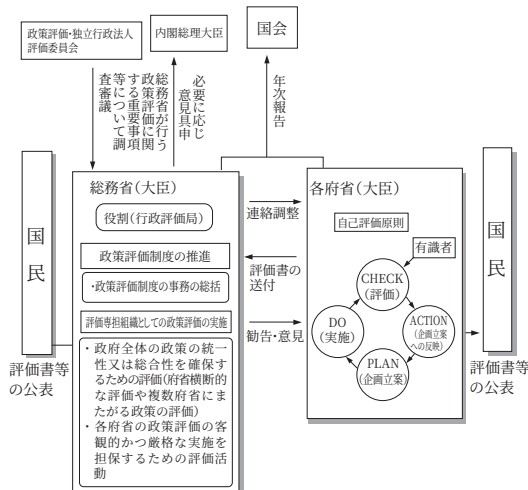
徹底、・複数府省が関係する政策（総務省が府省横断的に評価）、・政策評価の重点化・効率化、・規制の事前評価の早期義務付けに向けた取り組み、である。

政策評価法について、総務省は各府省、総務省および国会、国民という主要なステークホルダーズ間の関係を図2のように示しているが、政策評価に関する基本方針の改定が行われても、この基本的な構図は変わっていない¹⁷。

②経済財政諮問会議と政策評価との連携強化による関係図の変化

しかしながら、図2の構図に変化をもたらす決定が平成19年6月になされている。それは、経済財政諮問会議と総務省・各府省の政策評価に関する連携を強化する趣旨の閣議決定である（平成19年（2007年）6月19日）。すなわち、経済財政諮問会議の「基本方針2007」第3章「21世紀型行財政システムの構築、3. 予算制度改革」に記されたもので、政策評価を予算の効率化に適切に反映することが謳われている。具体的には、経済財政諮問会議は、総務大臣の「政策評価・独立行政法人評価委員会」に係る意見をもとに重要対象分野を提示する。総務大臣はこれを踏まえて政策評価を実施するというものである。

図2 政策評価結果の流れと総務省の役割



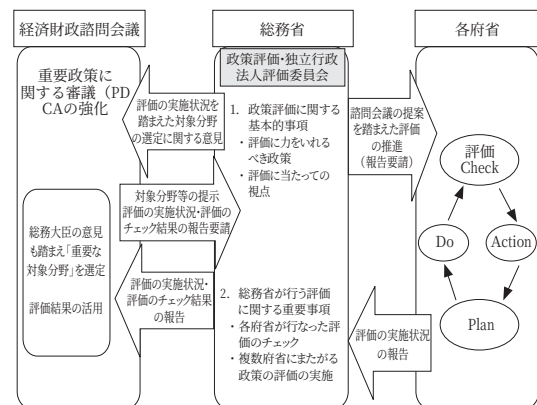
(出所) 総務省行政評価局HP

図3は平成19年5月9日総務省が菅議員に提出した資料（政策評価に係る諮問会議と総務省・各府省との連携）を参考に経済財政諮問会議と政策評価にかかる関係省庁との関係を示したものである。従来の総務省と各府省を中心とした関係に（図2）、経済財政諮問会議が政策評価の重点分野の提示と評価結果の被報告者として具体的に連携し関与することによって、政策評価と府省の上位にある政策決定プロセスとの間により密接な関係を築かれるかたちになっている。

また、政策評価に関する基本方針（平成19年3月30日変更版）III 1(2)において、政策評価の対象として、「施政方針演説等内閣の基本的な方針等により重点的に取り組むべきこととされた行政分野」とされており、政権が掲げた政策が政策評価の重点分野であることを明示している。

以上から、政策評価制度はその改変のプロセスの中で、政策評価対象の重点分野として施政方針演説など政治が掲げた政策を挙げ、また、内閣総理大臣の諮問機関である経済財政諮問会議との連携強化が行われた。このことにより、各府省の上位にある政策決定プロセスとの具体的なつながりが作られ、政策評価結果の政策への反映機会はより確保されてきたと捉えることができる。

図3 政策評価にかかる経済財政諮問会議との各府省との連携

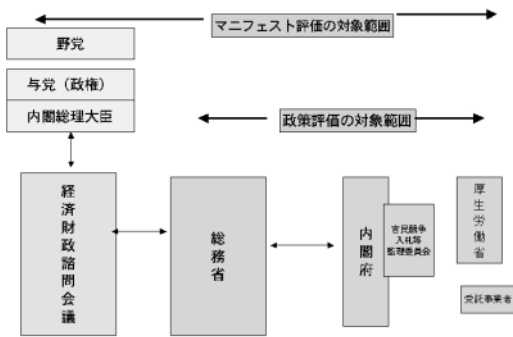


(出所) 総務省資料（平成19年5月9日）を参考に筆者作成

[illegible]

政策評価結果を次期計画立案へ活用し、PDCAサイクルを強化することによって国民へのアカウンタビリティを高めようとするのならば、行政機関が行う政策評価の上位に位置し政策決定について影響力を有する機関をも評価対象に含めることが必要であるのではないか。つまり、有権者と政治との約束を伴う政治決定の領域と現行の政策評価の2つの領域をつなぐ体系の構築という解が導かれる。政権構想や諮問会議の議論も評価対象に含め、その結果を以て、有権者に判断材料を提示しようとするマニフェスト評価はこの点でヒントを提示していると考える。図5は政策評価とマニフェスト評価の対象範囲を示したものであるが、現行の政策評価制度では対象は総務省に留ま

図5 政策評価とマニフェスト評価の対象範囲



(出所) 筆者作成

るが、マニフェスト評価の場合には経済財政諮問会議、政権が掲げた構想やマニフェストも対象になることを示している。

(3) 4つのレベルのPDCAとマニフェスト評価

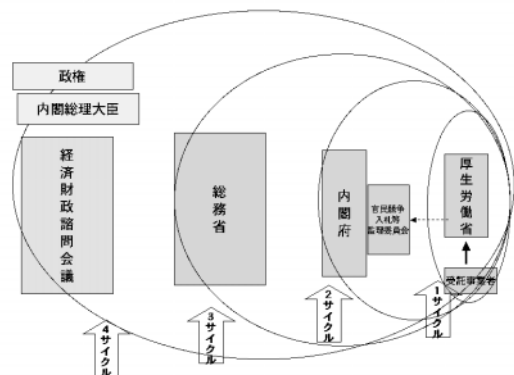
ここで提起したいのは、現行の政策評価かマニフェスト評価かという対峙や二者択一の問題ではない。政策評価の主軸であるPDCAサイクルと国民のアカウンタビリティを達成するためには、政策評価の対象をどのレベルまで引き上げる必要があるかという問題である。図6は市場化テスト制度（ハローワーク事業）を題材にマニフェスト評価の対象範囲にPDCAサイクルを付記したものである。ハローワークの受注業者と省庁を囲む円は市場化テスト制度下の個別事業レベルのPDCAサイクルであるので第1サイクルとした。官民競争入札等監視委員会は複数省庁が実施する市場化テスト事業全体を調整・監督する立場にあること、また未だ実施されていないが市場化テスト制度全体の評価は内閣総理大臣が行うことから暫定的に評価者として内閣府をここに位置づけた。そして、両者が市場化テスト全体のPDCAサイクルを担うことになるため第2サイクルとした。また、総務省は各府省から政策評価結果報告を受け、政策評価書の取りまとめや精査を担う役割があることからこのレベルを第3サイクルとした。内閣総理大臣が委員長をつとめる経済財政諮問会議は、政治レベルの政策決定プロセスに属するものとして、

第4サイクルに位置づけた。

前述のとおり、政策評価法に基づく評価の対象は総務省までであるので第3サイクルに留まることになる。他方、マニフェスト評価の場合には第4サイクルまでを評価の対象にする。また、ここでは、評価結果あるいは政策の実施結果から示唆されている内容と異なる方向の決定をする場合には、その理由も含め明確な説明を国民に行うことがアカウンタビリティ基準上、重要であるとされている。そして、このアカウンタビリティ基準に基づくマニフェスト評価結果が有権者に提示され、有権者がそれを投票の際の判断材料にし、政治がそのことを意識するようになれば、第4サイクルも機能するようになってくる。換言すれば第4サイクルの存在によって、政策の実施結果やその評価結果を逸脱して不鮮明なたちで政策を決定する行為を抑制する方向に向けられるのではないだろうか。

ただし、第4サイクルはそれ単独で機能できるわけではない。個別事業レベル（第1レベル）、制度運営に全体にかかるレベル（第2、第3レベル）のそれぞれの視点からの評価結果のフィードバックがあることによって第4レベルでの政策判断が可能になる。したがって、第1、第2、第3の各レベルのPDCAサイクルが機能してこそ、第4レベルのPDCAサイクルが成立すると考えるのが適当である。政策評価法に基づく制度設計上は第1から第3レベルのPDCAと総務省と各府省の間での

図6 4つのレベルでのPDCAサイクル



(出所) 筆者作成

評価情報のフィードバックは確保されているので、第4レベルにつなぐ方法をどう構築してゆくかということになる。また、上位のレベルでPDCAが機能しないと、下位のレベルも機能しなくなり負の連鎖を起こしうる。したがって、第4レベルにおいても、それがPDCAに沿って運営されることが課題になる。第4のレベルを対象として、そのPDCAサイクルが機能することを重視するマニフェスト評価アプローチは、第4のレベルにつなぐための解として、ひとつの方向を提示していると考えることができる。

おわりに

本論はハローワーク市場化テスト事業のPDCAサイクルに着目し、マニフェスト評価アプローチを用いて問題の所在を明らかにしようとした。この分析を通して、現行の政策評価の課題がより鮮明になった。すなわち、府省の上位にある政治的な決定によって、政策評価結果が無視されることがありえるということである。また、そうであれば政策評価のバックボーンであるPDCAサイクルが機能しないことになり、政策評価そのものの存在意義が問われることになる。

この問題の根底には政策のPDCAサイクルが国民に開かれていないという点にあると思われる。たしかに、国民への説明責任を政策評価の目的に謳っているが、ここで掲げられている政策（plan）について国民が不満に思っても投票のように、直接影響を及ぼすことはできないので、その立場は傍観者である。その意味で国民にとって閉じたサイクルなのである。だが、この閉じた政策のPDCAサイクルは、それ自体が独立して存在するサイクルではなく、その上位に有権者との約束による政策決定のサイクルがあり、それらがつながることで、政策決定のPDCAサイクルが機能することになる。だが、この上位の政策決定プロセスが考慮されないために、現行政策評価のPDCAサイクルが容易に壊されるという皮肉な結果を招いている。

2001年の政策評価法施行以来、政策評価には多大な予算と人的エネルギーが注がれてきた。だが、

この問題を解決しなければ、早晚、政策評価の意味が問われることになるのではないか。

では、政策評価をどのようにデザインし直せばよいのか。その際に重要なのは現行の政策評価が政策の決定とその実行の限られた一プロセスの評価を行っており、政治の判断がそのプロセスに大きな影響力を与えることを直視することである。その上で現行の政策評価に対して、有権者の参加を伴う政治のプロセスを如何に加えるかを考える段階に来ているように思われる。

こうした国民に「開かれた評価」が可能となるのは、政策決定が国民との約束を軸に行われるプロセスが機能する必要がある。日本の政治は現段階で政治が有権者に提示した公約を軸に回っているとは言い難いが、PDCAサイクルの重視が言われ、各官庁が政策評価を政策の実行に合わせて行っている中で、こうした「開かれた評価」が現状の政策評価へ強い示唆を与えているように考える。さらに個別の政策が上位の国家目標のもとで体系性が問われ、官自体の効率性が政策課題に挙げられる現状において省庁別評価に限界も見え始めている。

本稿で分析した市場化テストを巡る評価はその問題の構造を明らかにするものである。

有権者との約束に基づいて政策の組み立てが行われ、政府がその約束を実践する立場で政策を実行する場合は、政策の決定と実行、さらにその評価が、政党が有権者に提起する公約（マニフェスト）を軸にサイクルとして回る必要がある。そうした政策実行のプロセスはマニフェストを軸とした政治を実現させている英国などと比べわが国では完全に実現したわけではなく、その実現には機能別予算の導入など様々な課題がある。だが、すでに政党は選挙の際にマニフェストを公表するようになり、政府としてもその実行に責任が問われるようになっている。

政治主導での政策の決定や実現の動きが広がる中で政策評価の仕組みをどう組み合わせるかが本論の目的でもあったが、それは同時に政策評価に問われたアカウンタビリティの課題に答えを用意することでもある。各省庁別の評価を前提にその上位で政府全体の約束の達成を政府が自己評価し、有権者に評価が公表され、政府外の様々な

外部評価と評価の競争が行われることで、政策評価の内容をより国民に開くことができる。

いわば、PDCAのサイクルの中に有権者の評価（投票）も組み込まれることで、政策評価の第一に掲げられた目標であるアカウンタビリティ、すなわち国民への説明責任を果たすことにもなると考える。

注記

- 1 モデル事業については官民競争入札等管理委員会の文書で「民間から寄せられた提案のうち、モデル事業にならなかったものも、引き続き、制度の本格導入に向け、市場化テスト事業の対象とすることにつき検討を行う」とある。また規制改革・民間開放推進3ヵ年計画において、「モデル事業を実施し、また本格的導入に向けた制度を検討する上での指針として」と記されており、モデル事業が本格導入のための試行として位置づけられていることは明らかである。
- 2 競争の導入による公共サービス改革に関する法律（市場化テスト法）平成21年法律第51号
- 3 2007年5月時点において官民競争入札等監理委員会事務局および所轄庁である厚生労働省に問い合わせたが、監理委員会より公開されていないとの回答があった。
- 4 官民競争入札等監理委員会 2007年6月26日の議事録より
- 5 厚生労働省は北海道旭川市周辺、青森市周辺、高知周辺、福岡県飯塚市周辺、長崎県佐世保市周辺の5地域に民間開放する方針を決定したが、2007年3月の入札時期において応札者が出たのが2ヶ所のみであり、その後再入札をかけて3ヶ所が決定された。
- 6 官民競争入札等監理委員会事務局へのインタビューから（2007年5月）。
- 7 ILO第88号条約においては、職業安定組織は国の指揮監督の下にある職業安定機関の全国的体系で構成され、無料で、かつその職員を公務員としなければならない旨が規定されていること
- 8 「キャリア交流プラザ事業」「若者版キャリア交流プラザ事業」「求人開拓事業」
- 9 21世紀民間臨調で発表した9団体の中で、最も重厚なマニフェスト評価を実施したと東京新聞において論評されている（2007.7.2）。
- 10 ・年金・社会保険庁改革、・医療、・少子化対策、・NPO・公益法人、・市場化テスト、・構造改革特区、・格差・再チャレンジ、・外交・安全保障、・国と地方、・地球環境、・財政政権、・経済政策、・農業・食糧、・治安、・公務員制度改革、・教育改革、・政治とカネ、・憲法改正・国民投票の21の政策
- 11 安倍政権は選挙を経て誕生した政権ではないので、安倍政権の実績を評価する時点（2007.6）において、それに対応するマニフェストは存在していない。そこで、評価対象として政権構想にあたる「美しい日本」や所信表明演説などを政策目標として捉えた。
- 12 ちなみに7月に提示した自民党マニフェストにも官民パートナーシップに関する言説はなくなっている。
- 13 第7回経済財政諮問会議 議事録より
- 14 官民競争入札等監理委員会事務局へのインタビューから（2007年5月）。
- 15 ヒアリング対象者は、地方で求人開拓事業とキャリア交流プラザ事業の双方を2年以上にわたり受注した企業の社長と担当者である。
ヒアリング項目は次のとおりである。
・参入の動機と狙い、・入札方法、・予定価格、・委託契約金と実費（採算性も含めて）、仕様書、・引き継ぎ、・業務運営の実際、・困難な対象地域および対象者への対処方針と対応方法、・評価結果、・評価結果のフィードバック、・意見申し立て。
- 16 経済財政諮問会議がハローワークへの市場化テスト導入を急がせた背景には、直前にひかえた参議院選挙（2007.7）の問題があると思われる。すなわち、選挙での改革色を強調する意図が働いていたのではないか。
- 17 平成19年9月より、規制影響分析（RIA）が義務づけられることになった。これによって全ての規制は閣議決定前に規制がもたらす費用と便益を分析し提出することが定められる。その意味で政策の事前評価機能がより強化されたといえる。ただし、RIAの位置づけはあくまでも既存の政策評価の事前評価と位置づけられおり、図2の関係図に変化をもたらすものではない。

参考文献

飯尾潤（2003）『日本の選挙公約と書が帰国のマニフェスト：その対比から』自由民主党国家戦略本部第27回講演、2003年3月19日

大山礼子・藤森克彦（2004）『マニフェストで政治を育てる』、雅粒社

言論NPO（2003）『マニフェストと日本の論点』

言論ブログ・ブックレット（2007）『日本の政治を採点する 2007年参議院選の公約検証』、言論NPO

言論ブログ・ブックレット（2007）『安倍政権の通信簿 政権300日の実績を評価する』、言論NPO

言論ブログ・ブックレット（2007）『美しい国とは何か』、言論NPO

言論ブログ・ブックレット（2007）『安倍政権100日を評価する』、言論NPO

榊原秀訓・家田愛子・尾林芳匡（2006）『イギリスの市場化テストと日本の行政』、自治体研究社

曾根泰教（2004）「マニフェストの原点 選挙そのものも姿が変わり政権党の実行体制が強く要求される」、We believe 2004年1月号

田中弥生（2006）「構造調整借款20年のレビューからみる日本政府の政策と判断」、『日本評価研究』日本評価学会第6号、(特活)日本評価学会

田辺国昭（2005）「中央省庁における政策評価の現状と課題 期待と失望のスパイラルを超えて」、『政策形成支援のための政策評価 NIRA型政策評価モデルの提言 NIRA研究報告書0506』、第2章第1節、総合開発研究機構

内閣府公共サービス改革推進室（2006）『詳細 公共サービス改革法』、ぎょうせい

内閣府公共サービス改革推進室（2006）『よくわかる！ 公共サービス改革法入門』、ぎょうせい

『特集 市場化テストが官製市場を改革して生じるメリットは？』、『法律文化5』、2006vol.263、LEC 東京リーガルマインド

本間正明監修（2005）『市場化テスト 官民競争の到来』、NTT出版

「ニッポンの公共サービス 公務員」『週間東洋経済 7/14』、東洋経済新報社

八代尚宏編（2005）『「官製市場」改革』、日本経済新聞社

山谷清志（2006）『政策評価の実践とその課題 アカ

ウンタビリティのジレンマ』、萌書房

総務省行政評価局「政策評価に関する基本資料」、平成19年11月

「行政機関が行う政策評価に関する法律」（平成13年6月29日法律第86号、最終改定平成15年法律第23号）

「政策評価に関する基本方針」（平成17年12月16日、平成19年3月30日）

総務省報道資料「規制影響分析（RIA）の試行的実施状況」（平成19年10月26日）

総務省行政評価局「政策評価Q&A」平成19年11月

総務省行政評価局 政策評価 2007.8.30 http://www.soumu.go.jp/hyouka/QandA/pdf/060710_sh_qa_2-1.pdf

総務省報道資料 平成19年11月12日「政策評価の重点対象分野の選定等について（政策評価・独立行政法人評価委員会の答申）」

総務省「政策評価の機能の発揮に向けて ー重要政策に関する現状と課題ー」平成19年5月9日菅議員提出資料

「美しい国、日本」政権構想 2006.9.26 <http://newtop.s-abo.or.jp/front-page/profile.html>

「第165回国会における安倍内閣総理大臣所信表明演説」2006.9.29（首相官邸）<http://www.kantei.go.jp/jp/singi/utukusii/abespeech/165.html>

「第166回国会における安倍内閣総理大臣施政方針演説」2007.1.26（首相官邸）<http://www.kantei.go.jp/jp/abepphoto/2007/01/26sisei.html>

「日本経済の進路と戦略について」2007.1.25閣議決定（首相官邸）<http://www.kantei.go.jp/jp/singi/keizai/kakugui/070125kettei.pdf>

「行政改革会議 平成8年11月発足」1996.11（首相官邸）<http://www.kantei.go.jp/jp/gyokaku/1203dai1.html>

「行政改革会議 労働省説明資料」平成9年5月7日 1997.5.7（首相官邸）<http://www.kantei.go.jp/jp/gyokaku/roudo507-1.html>

「総合規制改革会議 平成13年4月発足 「民間参入の拡大による官製市場の見直し」」（首相官邸）2001.4 <http://www.kantei.go.jp/jp/singi/kisei/index.html>

「規制改革・民間開放推進会議 平成16年4月」2004.9（内閣府）http://www8.cao.go.jp/kisei-kaikaku/old/minutes/commission/2004/09/item04_09_03.pdf

「規制改革・民間開放推進会議第1次答申」2004.3.22（内閣府）http://www8.cao.go.jp/kisei-kaikaku/old/market/2004/0322_02/market040322_02_03-01.pdf

「規制改革・民間開放推進会議第2次答申」2005.12.21
(内閣府) http://www8.cao.go.jp/kisei-kaikaku/old/market/2005/1221/item051221_01.pdf

「ハローワークとILO条約に関する懇談会の開催について」2006.12.19 (内閣府 公共サービス改革推進室)
<http://www5.cao.go.jp/koukyo/ilo/files/061219.pdf>

「ハローワークとILO条約に関する懇談会 報告書について」2007.4.6 (内閣府 公共サービス改革推進室)
<http://www5.cao.go.jp/koukyo/ilo/files/report1.pdf>

「改革と展望 2002年度改定」2003.1 (厚生労働省)
<http://www.mhlw.go.jp/shingi/2003/01/dl/s0122-3e.pdf>

「構造改革と経済財政の中期展望」平成14年1月25日
(2002.1.25) (総務省) http://www.soumu.go.jp/czaisei/czaisei_seido/ichiran03_e.pdf

「構造改革と経済財政の中期展望」2005年度改訂版
(経済財政諮問会議提出) <http://www1.odn.ne.jp/hnakashima/cyuki2005.pdf>

「経済財政運営と構造改革に関する基本方針について 2006」2006.7.7 (経済財政諮問会議) <http://www.kantei.go.jp/jp/singi/keizai/kakugi/060707honebuto.pdf>

「基本方針2006の総点検について」平成19年4月6日
(2007.4.6) (経済財政諮問会議) http://www.keizai-shimon.go.jp/minutes/2007/0406/item11_1.html

「経済財政改革の基本方針について 2007」(2007.6.20)
(経済財政諮問会議) <http://www.keizai-shimon.go.jp/cabinet/2007/decision070620.pdf>

「規制改革・民間開放推進3カ年計画(改定)」平成17年3月25日閣議決定(2005.3.25) (内閣府 規制改革・民間開放推進会議) http://www.stat.go.jp/info/kenkyu/minkan/pdf/sanko1_2.pdf

「『市場化テスト(モデル事業)』の適切な実施に向けて」平成17年3月14日(2005.3.14) (内閣府 規制改革・民間開放推進会議) <http://www8.cao.go.jp/kisei-kaikaku/old/market/index.html>

「市場化テストに関する海外事例調査報告書」平成17年10月(2005.10) (内閣府 規制改革会議・民間開放推進会議 市場化テスト・ワーキング・グループ) <http://www8.cao.go.jp/kisei-kaikaku/old/market/2006/item060607.pdf>

「『市場化テスト(官民競争入札制度)』に関する民間提案の募集の受付状況について」2005.11.19 (内閣府 規制改革・民間開放推進室会議) <http://www8.cao.go.jp/kisei-kaikaku/old/market/2004/1119/market>

041119_01.pdf

「競争の導入による公共サービスに関する法律」(平成十八年六月二日歩率第五十一号)

「公共サービス改革基本方針 平成18年12月」(2006.12)
(内閣府公共サービス改革推進室) <http://www5.cao.go.jp/koukyo/kihon/200612/kihon.pdf>

「競争の導入による公共サービスの改革について 公共サービス改革法」2007.5 (内閣府公共サービス改革推進室) <http://www5.cao.go.jp/koukyo/kaisetsu/leaf/0.pdf>

「競争の導入による公共サービスの改革について 市場化テスト」2006.5 (内閣府公共サービス改革推進室)(ヒアリング先より入手)

「第21回官民競争入札等監理委員会議事」平成19年5月23日(内閣府官民競争入札等監理委員会) <http://www5.cao.go.jp/kanmin/kaisai/2007/523/523.html>

「市場化テストに関するスコアカード」(2007.3) (内閣府官民競争入札等監理委員会) <http://www5.cao.go.jp/kanmin/scorecard/scorecard.html>

「ハローワークの市場化テストの主な論点について」2007.7.30 (内閣府官民競争入札等監理委員会) <http://www5.cao.go.jp/kanmin/kaisai/bukai/hellowork/2007/730/730.html>

「官民競争入札等監理委員会(資料)」2007.4.24 (厚生労働省) <http://www5.cao.go.jp/kanmin/kaisai/bukai/hellowork/2007/730/730.html>

「ハローワークの市場化テストについて(有識者議員提出資料)」2007.5.9 (経済財政諮問会議) <http://www.keizai-shimon.go.jp/minutes/2007/0509/item7.pdf>

「ハローワークについて」柳澤臨時議員提出資料 2007.5.9 (経済財政諮問会議) <http://www.keizai-shimon.go.jp/minutes/2007/0509/item6.pdf>

「平成19年第7回経済財政諮問会議訪問会議議事要旨」2007.4.6 (経済財政諮問会議) <http://www.keizai-shimon.go.jp/minutes/2007/0406/shimon-s.pdf>

「ハローワーク市場化テストの導入(民間議員提出資料)」2007.4.6 (経済財政諮問会議) <http://www.keizai-shimon.go.jp/minutes/2007/0406/item2.pdf>

「ハローワークとILO条約に関する懇談会」報告書 2007.4.6 (経済財政諮問会議) <http://www.keizai-shimon.go.jp/minutes/2007/0406/report.html>

「平成19年第3回経済財政諮問会議議事要旨」2007.2.16 (経済財政諮問会議) <http://www.keizai-shimon.go.jp/>

minutes/2007/0216/shimon-s.pdf

第3回経済財政諮問会議資料「市場化テストの今後の進め方について」2007.2.16（官民競争入札等監理委員会）<http://www.keizai-shimon.go.jp/minutes/2007/0216/item6.pdf>

「ハローワークの市場化テスト導入について」2006.11.30（経済財政諮問会議）<http://www.keizai-shimon.go.jp/minutes/2006/1130/item5.pdf>

「国民年金保険料の収納業務（市場化テストモデル事業）に係る評価について」平成19年2月（社会保険庁運営部年金保健課）<http://www.sia.go.jp/top/kaikaku/shijoka/shijoka070208.pdf>

「平成17年度市場化テストモデル事業（キャリア交流プラザ事業、若年者版キャリア交流プラザ事業及び求人開拓事業）に係る実績評価について」平成19年6月1日（厚生労働省）<http://www.mhlw.go.jp/houdou/>

2007/06/h0601-3.html

「各府省が実施した政策評価についての審査」平成14年度総括（総務省）http://www.soumu.go.jp/s-news/2002/021205_3_01.html

「神吉伸之の選挙コラム」2007 <http://www.election.co.jp/column/index2.html>

朝日新聞「保険料徴収コスト、民間委託で3～5割減 市場化テスト」、朝刊2007.2.8

朝日新聞「市場化テスト黄信号 求人開拓事業3地域で落札なし」、朝刊2007.3.20

朝日新聞「政策目標を就業率の向上に＝労働市場改革で大田大臣」、朝刊2007.4.6

朝日新聞「ハローワーク官民競争」、2007.5.9

東京新聞「参院選2007」、朝刊2007.7.2

（2008.1.23受理）

Examining the Limitation of the Current Policy Evaluation System in Japan: The Case of Policy Management Problems of the Market Testing System Executed at HELLO WORK

Yayoi Tanaka

National Institution for Academic Degrees
and University Evaluation
ytanaka@niad.ac.jp

Yasushi Kudo

Genron NPO
y-npo@mta.biglobe.ne.jp

Abstract

The purpose of the paper is to point out the problem of the current policy evaluation system with case analysis of the market testing system that was executed by HELLO WORK, the employment bureau under the Ministry of Health, Labor and Welfare. The marketing test system provides sufficient conditions to manage project or policy cycle since the ex ante and the ex post evaluations are set in within the system. However, the policy management system of the marketing system executed by HELLO WORK did not function well because of the strong influence of decisions by the Council on Economic and Fiscal Policy which is positioned in the upper level of the Ministry of Health, Labor and Welfare. The current policy evaluation however could not disclose this problem. This paper clarifies this problem with the new approach of manifest policy evaluation that includes the policy of the ruling party and its execution process as evaluation targets. The structural differences between the current policy evaluation system and manifest policy evaluation are then analyzed in order to address the problems of the current policy evaluation system and to recommend the direction of its solution.

Keywords

Policy Evaluation, Manifest Evaluation, Market Testing System,
HELLO WORK(employment bureau), policy management

【研究論文】

自治体の評価能力向上研究 — 「分かりやすい」評価説明のために —

高千穂 安長

玉川大学

takatiho@bus.tamagawa.ac.jp

要 約

自治体の政策評価においては評価を正しく行うとともに、正しく伝えるという努力も求められる。

本稿では、評価担当者の評価への動機付けから、所属部局など個々の組織、自治体、市民など外部との評価活動に至る因果関係を明示した後、川崎市の実例をもとに評価結果の「分かりやすさ」確保のための対応の現状について研究した。

その結果、「分かりやすく」ない評価報告がなされる原因は、「知」、「動機付け」が無いためではなく、評価者の視点の転換、時間に追われる中での効果的評価実施などがなされていないことに起因することが明らかとなった。

キーワード

分かりやすさ、説明責任、評価の視点

1. はじめに

現在、政策評価は多くの自治体で導入され、その成果についても肯定的な研究成果（横山 2006）が出されている。しかし、「行政の説明責任が国民本位の効率的で質の高い行政の実現、国民的視点に立った成果重視の行政の転換という評価目標も国民の行政の説明責任に収斂するため、説明責任が非常に問題になる」（山谷 2006）ことから、「政策評価結果を本来積極的に活用すべき側の国民にあまり影響を与えていない」（金城 2002）という意見があるのは問題となる。総務省も「政策評価制度に関する見直しの論点整理⑧評価書の簡明さの確保」（2004年1221日）で、各府省が行っ

た政策評価の結果を国民に分かりやすく伝えることが重要としており、そのためにはどのような方が考えられるかを問題としている。

しかし、現実問題として、政策評価の導入、導入後の評価理論については多くの研究成果があるが、国民に対する説明（アカウンタビリティ）についての研究はあまりされていない。

本稿では、政策評価に当たって重要なテーマであるが、あまり研究が行われていない評価結果の説明について、川崎市¹が行っている施策評価²に対する取り組みをもとに研究を行う。

2. 評価サイクルと仕組み

評価は通常、P-D-C-Aのサイクルに基づき、「目的—手段」を明確にし、その進捗状況に基づいて行う。実際に評価を行う段階では、評価担当者の能力、意欲とその所属組織の文化、能力、および外部環境の影響を受ける。

評価結果を「分かりやすく」公表するためには、これらがそれぞれ評価志向的な体制になっていることが重要となる。

(1) 「分かりやすい」評価の関係

図1は、一般論として評価作業を行う者が「分かりやすい」評価を行うための個人、組織、外部の関係を示している。

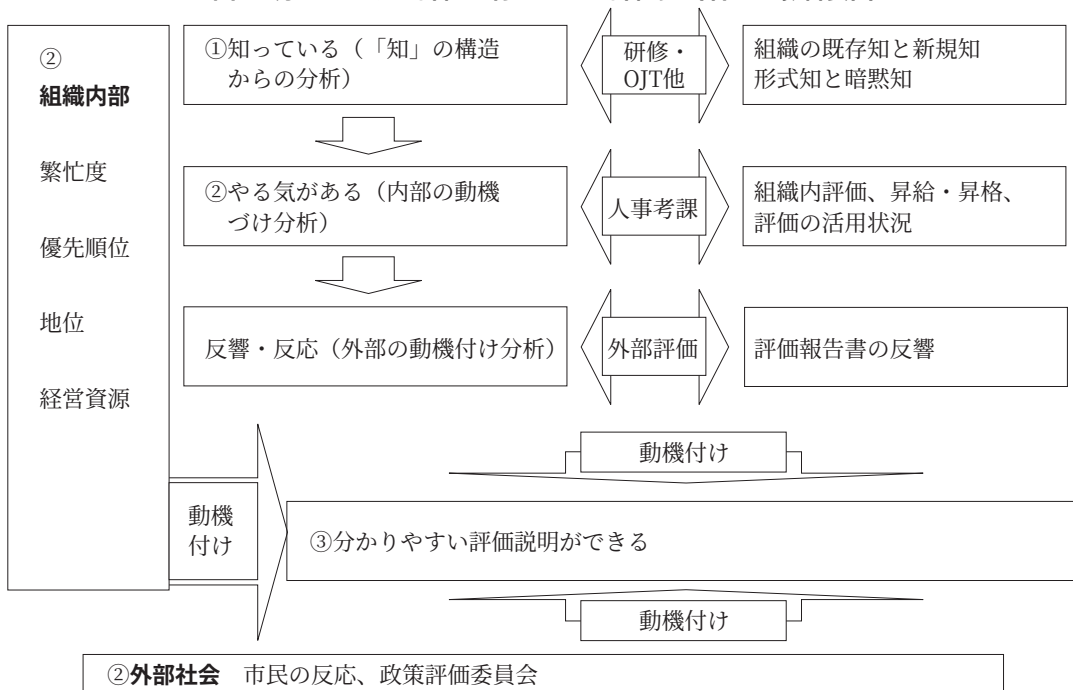
①では、評価を行うには、まず、個人が評価実施についての必要な「知」の入手の必要性を書いている。「知」と一言言われるが、「知の構造」として先ず「既存知」と「新規知」に区分する必要がある。「既存知」は組織文化としてすでに認

知され分かっていることであり、多くは「マニュアル」などに文字情報として収録されている。「新規知」は、文字通り新しい知識であり、組織、個人として新規に受け入れる必要がある。評価に関する新規の理論などがこれに該当する。新規知については、既存知と往々にして対立するケースが多い。このため、混乱や葛藤が生じる（上笹2007）。

次いで「知の構造」は「形式知」と「暗黙知」に分類できる³。「形式知」はマニュアルや研修会などで目や耳を通して獲得できるが、「暗黙知」については、実際に評価を行わないと理解できない。「知」について、既存知のうち形式知はしっかりと継承し、暗黙知をOJTなどにより伝播することが求められる。また、必要な新規知の取り込みに当たっては混乱、葛藤を鎮めることが必要となる。

②では、「動機付け」の必要性を書いている。ハーズバーグの二要因説⁴からは、評価担当者が「知」を得た後、分かりやすい報告書作成について十分な努力を行う仕組み作りの上で有益な示唆

図1 分かりやすい評価を行うための評価担当者への影響要因



を得る。

自治体の原課評価担当者は、行革推進中であることから、評価業務だけではなく、多くの実務を兼務している。その中で評価実務の重要性を理解し、分かりやすい表現に時間を割くのは、実は相当に重い負担となる。限られた時間の中で納期を守るためには、可能な限り「効率良く評価を行い、文書化する」必要があるため、「既存の文書などをコピー＆ペーストで行う（計画に書かれていることをそのまま使う）」ことも行われる。このような状況下で評価担当者の積極的に評価に取り組む姿勢を生み出すには、組織内部では「つけたし」的な業務認識ではなく、「エース」が行う業務という位置づけが求められる。また、外部要因としては政策評価委員会、市民の意見のフィードバックによる「役立ち認識」が重要となる。

3. 分かりやすい評価報告とそのための体制

(1) 一般的な分かりやすい評価報告

市民にとって分かりやすい評価報告となるには、基本的には評価の原型である「目的⇒手段」が明確であることが必要となる。

一般論としては、次の視点に基づいて評価結果が記載される必要があると考えられる。（表1）

(2) 自治体政策評価報告書の「分かりやすさ」の体制

自治体政策評価報告の例として、川崎市の政策評価結果の分かりやすさを取り上げ検証する。

川崎市の政策評価は、図2の通り先ず実施部署が内部評価を行う。この結果を政策評価委員会がテーマ毎に検討を行い、公平性を担保する。さらに、そのアドバイスをもとに修正を行い、その結果について公示をし、得られた市民意見を反映する手順になっている。このように、市民に「評価情報と代替案をなるべく広く透明に顧客、利害関係者と一般市民へ広報する」（上野 2004）というアカウンタビリティの目的を達せられる体制が作られている。

本稿で問題とするのは図2の①の段階であり、原課の評価報告が「どの程度の分かりやすさ」かを問題とし、問題がある場合は「どのような問題なのか」を示し、その対策について既述の先行研究を元に対応を検討する。

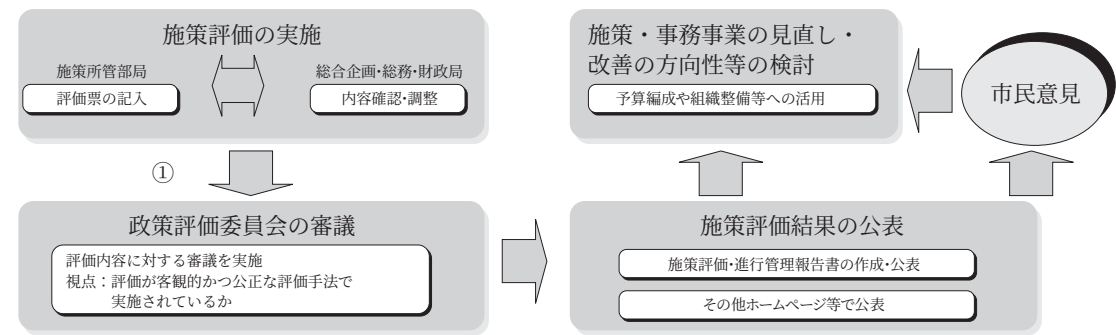
4. 「分かりやすさ」についての現状

ここでは、川崎市の実例に即して状況を明らかにする。

表1 「分かりやすい」評価の視点

項目	内容
施策目標	1. 目標の明確性 <視点> 市民が目標を容易にイメージできる記述になっているか 2. 目標の妥当性 <視点> 解決すべき課題と目標の方向性の関係が妥当か
成果の説明	1. 説明の妥当性 <視点> 施策の目標を踏まえた成果の説明となっているか 1. 説明の客観性 <視点> 客観的な分析か 3. 説明の分かりやすさ <視点> 分かりやすい成果の説明か

図2 川崎市の政策評価体制



(出所) 川崎市

表2 「分かりやすさ」調査

単位：％

得点/ 視点	施策目標記述		成果説明記述			合計
	目標明確性	目標妥当性	説明の妥当性	説明の客観性	説明の分かりやすさ	
4	11.7	13.1	5.6	2.8	5.6	7.8
3	57.5	53.7	58.4	56.1	65.4	58.2
2	27.6	29.4	29.9	36.0	24.3	29.4
1	3.3	3.7	6.1	5.1	4.7	4.6
合計	100	100	100	100	100	100

(出所) 川崎市

表3 平成17-18年度施策評価の内部評価結果

評価区分	内容	施策課題構成比(％)	
		H17年度	H18年度
A	成果があり、課題解決の阻害要因なし	60.0	58.8
B	成果はあるが、課題解決の阻害要因あり	37.3	38.8
C	成果が上がっていない	0	0
Q	現時点での評価困難	2.7	2.4
合計		100	100

(出所) 川崎市

(1) 川崎市の評価結果の分かりやすさ

表2の通り、評価結果の説明については、概ね妥当（得点3以上）が66％あり、2/3が問題なしとなっている。しかし、1/3は「分かりやすく」無いということであり、改善を要することが分かる。「分かりやすい」説明が阻害される要因について、次の仮説に基づき、検討を行う。

(2) 「分かりやすさ」を阻害する要因についての仮説

仮説① 施策の進捗が良くないために説明にムリがある

「分かりやすさ」を阻害する要因のひとつとして、施策の進捗状況が良くないために、分かりに

表4 部署別「分かりやすさ」の分析

	目標明確性	目標妥当性	説明妥当性	説明客観性	説明の分かりやすさ
局1	0.27	0.27	0.64	0.64	0.73
局2	1	1	1	0.5	1
局3	0.5	0.5	0.5	0	0.5
局4	0.25	0.63	0.75	0.25	0.5
局5	0.47	0.80	0.40	0.27	0.27
局6	0.5	0.5	0.5	0.25	0
局7	0	0	0.13	0.13	0.13
局8	0.5	0.41	0.55	0.77	0.27
局9	0.38	0.29	0.62	0.81	0.57
局10	0	0	0	0	1
局11	0	0.44	0.33	0.67	0.33

注：0=問題ない、1=問題あり。数値が低いほど問題はない。

局2のサンプル数は2、局10のサンプル数は1。

（出所）川崎市政策評価委員会

表5 項目別「分かりやすさ」

目標明確性	目標妥当性	説明妥当性	説明客観性	説明の分かりやすさ
0.40	0.43	0.51	0.52	0.41

（出所）川崎市政策評価委員会

くく書くことも考えられる。これについては、川崎市の実施している施策は、表3の通り、概ね良い進捗を示しており、問題はない。従ってこの仮説は妥当ではない。

仮説② 組織（特定部署）に問題がある

表4の通り、サンプル数が少ない局を除いても、部署毎の評価に対するばらつきはある。従って、特定部署が評価報告において、分かりにくくなっていると考えてよい。

本仮説は妥当と考えられる。

仮説③ 説明項目によって評価説明の難度が変わり、「分かりやすさ」が阻害される。

表5の通り、説明の妥当性、説明の客観性がや

や難しく、目標の明確性は比較的容易ということが読み取れる。

5. 「分かりやすさ」が阻害される具体例

評価部署により評価報告が分かりにくくなることはあり、評価項目により評価説明の「分かりやすさ」に差異がでることも判明した。では、どのように「分かりやすさ」が阻害されているのだろうか。表1の項目に従って具体例を検証する。

川崎市は施策評価報告のフォーマットを決めており、項目別に記述を行えば報告として完成するようになっている。「分かりやすさ」の問題を考えるには先ず「フォーマットの妥当性」を考える必要があるが、各項目については三重県などの例を参照して決めており、大差ないため、ここでは、

「記述する内容」に焦点を当て、表1の項目に従って事例をレビューする。なお、『『確かな学力』などあいまいな用語使用』、「跨道橋、インキュベータなどの専門用語の使用」、「カタカナ言葉の多用」、「『PFI』など略字の説明無き使用」など問題が分かりやすいものは事例として例示しない。

以下では、〈施策課題〉、〈施策の目標〉、〈当該施策によって解決すべき課題〉、〈施策の概要〉は、原課の報告書の内容をそのまま掲載、〈問題〉、〈問題の背景仮説〉は筆者の見解を示している。

(1) 目標を明確にイメージできる記述になっているか

〈施策課題〉 安全な地域社会の確立
 〈施策の目標〉 市民の生命や健やかな健康を守るため、地域で発生する犯罪、交通事故の防止に向けた安全な地域社会の確立を目指します。
 〈問題〉 目標に数値や改善状況の例示など具体的な記述がなく、「今年どうなったのか」などが分からない。
 〈問題の背景仮説〉 計画時の課目標をそのまま使用し、評価報告をする意識が無い。

(2) 「解決すべき課題」と目標の方向性との関係が妥当か

〈施策課題〉 救急救命士の養成と高度な救急体制の整備
 〈施策の目標〉 ①救急隊の救急救命士常時乗車体制を確立し、かつ病院実習や症例検討会を通じて救急救命士の知識・技術向上を図ることにより、高度な救急サービスを市民に提供します。②応急手当ができる市民救命士を、1世帯1人を目標に養成し、併せて救急車の適正利用についても普及啓発を行います。
 〈当該施策によって解決すべき課題〉 ①2006年の本市の救急出動件数は前年比-774件の58,862件で22年ぶりに減少したが、相変わらず救急搬送のうち約60%が軽症者であることから、より多くの市民が応急手当の知識や技能を修得する機会を増やすことにより、適正な救急車の利用について普及啓発し、真に救急車を必要としている傷病者の利用に支障が出ないようにする必要があります。②救急救命士の医療行為が拡大しその重要性が高まっ

ていることから、より高度な救命措置のできる救急救命士の常時乗車を早期に実現するために救急救命士を計画的に養成し、救急サービスを充実させる必要があります。

〈施策の概要〉 ①救急要請が増加する中、救急サービスの維持・向上を図るため、救急隊の到着時間の均等化を図り適正配置を行います。②高度な救命措置ができる救急救命士の救急車への常時乗車体制を早期に実現するために、救急救命士の養成を計画的に進め、気管挿管や薬剤投与など、より高度な救命措置ができる救急救命士を養成し、高度な救急サービスを市民に提供します。③救急隊到着前に応急手当が実施できる市民（市民救命士）を、1世帯1人を目標に養成します。あわせて、職員研修等の機会を捉え、全ての市職員を市民救命士として養成します。

〈問題〉 目標と施策（いわゆる目的一手段）の区分けができていない。課題—目標—施策とも、極端に言えば全て救急体制の高度化を図ることを言っているともとれる。

〈問題の背景仮説〉 投入=産出（アウトプット）=評価対象という意識はあるが、効果（アウトカム）を把握し、評価するという意識が乏しい（またはデータ収集体制が無い）。言い換えれば市民は何を知りたいのかという意識が乏しい。

(3) 「施策の目標」を踏まえた「成果の説明」になっているか

〈施策の目標〉 市内から排出される一般廃棄物を、年間を通して適正かつ安定的に処理し、快適な生活環境の保全を目指します。
 〈解決すべき課題に関する当該年度の主な成果〉
 ①廃棄物処理施設の性能を維持するために、法令等で定められた各種の検査及び点検整備業務を実施し、排ガス等の環境基準を遵守した適性かつ安定的な処理・処分を行いました。
 ②処理センターでは環境マネジメントシステムにより、日常的に適正な維持管理を実施し環境対策に万全を期すとともに、ISO14001（国際標準化機構が定める環境マネジメントシステムの国際規格）の定期審査に合格し、認証を継続しています。

〈問題〉 目標に数値や改善状況の例示など具体的な記述がなく、「今年どうなったのか」などが分からない。

〈問題の背景仮説〉 計画時の課目標をそのま

ま使用し、評価報告をしているという意識が無い。

(4) 客観的な分析か

〈施策の目標〉ごみの発生・排出抑制、資源物の分別収集の拡充など、循環型社会の構築を目指して3Rを推進します。

〈解決すべき課題に関する当該年度の主な成果〉

①分別収集品目の拡大を可能とする新たな収集体制に向けた検討を行い、2007年度からの作業体制を確立しました。②川崎・幸区においてミックスペーパーの分別収集モデル事業を11月から開始するとともに、2007年度のモデル事業対象地域の拡大及び収集業務の民間委託化に向けた準備を行いました。

(注) 参考指標欄に空き缶・ペットボトル収集量 計画策定時11,078t、2005年度11,895tの記載あり。

〈問題〉目標に数値や改善状況の例示など具体的な記述がなく、「今年どうなったのか」などが分からない。

〈問題の背景仮説〉評価フォーマットの各項目を記述することと、分かりやすさよりも二重記述を避ける意向の方が強い。読み手の読みやすさが前面に出ていない。

(5) 分かりやすい成果の説明か

〈施策の目標〉区役所を地域のまちづくり拠点として整備するとともに、関係する市民との協働を進め、地域のまちづくりに関する課題に総合的に取り組むことが目標です。

〈解決すべき課題に関する当該年度の主な成果〉

①区において、新たに路上喫煙防止対策の取組を開始するなど、身近な環境整備への取組を進めました。②区におけるまちづくり支援体制の強化に向けた検討を行い、平成19年度に向けて組織体制を整備することとしました。

③区におけるまちづくりに関する課題の解決に向けて、区における総合行政の推進に関する規則に基づく調整を行うなど、事業局と区役所との連携によって事業を実施しました。

〈問題〉書いてあることは分かるが、「何がどうなった、問題は何か」などの評価報告として必要な情報が伝えられていない。

〈問題の背景仮説〉「説明責任とは何か」、「何を市民に説明し、原課の必要性、有効性を伝えたい」という意識が無い。

6. 「分かりやすい」評価報告実施のために

既述の図1.③「分かりやすい評価説明」は、①「知」の取得、②内部・外部からの動機付けにより達成されるとしたが、これを4.の具体例、問題および問題の背景仮説を元にまとめると、次のように集約できる。

川崎市の場合、評価担当者は「知」や「動機付け」については部署により若干の差異はあるが、既に習得していると考えられる。具体例からは、評価担当者の視点として、①評価をしているという視点の欠如、すなわち、評価理論の基礎がわかっていないために、実施した事業の産出(アウトプット)の記述にとどまり、どのような成果が上がったか、持続的な仕組みができたかが示されていない、②市民の視点の欠如、すなわち、市民が知りたい情報はどのようなものかを考えていない、そのため、計画時の表現をそのまま使って十分としているなどが明確になった。

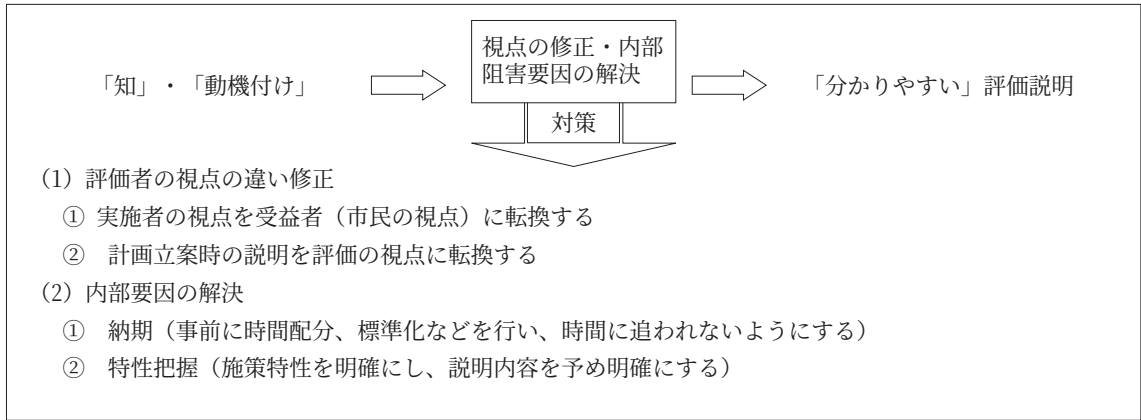
従って、課題と対策は、次頁の図3のように集約される。

7. おわりに

今回は「内部評価」結果についてメタ評価を行う形で、評価結果を正しく、わかりやすく伝えるための課題を明らかにした。今回の研究では、内部評価報告書からの知見をもとに、「分かりやすい」評価の方策について研究した。現実問題として複数年の事業計画などでは毎年の進捗評価にそぐわないものがあり、また、計画時に評価を意識していない計画策定から、評価担当者が適切な評価指標を見出せないという声もある。特定部署に評価に対する問題が存在する可能性が窺えたが、本稿では「特定の部署で多忙のあまり評価に十分な時間が割けない」⁵ということのみで対応した。これらについては次回の研究のテーマ⁶としたい。

必要なデータの使用を快諾いただいた川崎市総合企画局、川崎市政策評価委員会に感謝申し上げる。なお、文中の意見等は筆者の個人的意見であることを申し添える。

図3 課題と対策



（出所）筆者作成

注記

- 1 自治体を考える時、都道府県から市町村にまでまたがるため、どれを平均的な自治体と考えるかという問題が発生する。川崎市は政令指定都市であり、都道府県と市の二つの性格を有する。また、17の政令指定都市の中で、推定人口で8位と中位にあり、平均的な自治体とみなすことができると考えられる。
- 2 川崎市では「政策評価」と称して「施策評価」を行っている場面もある。本稿では、厳密な区分はしていない。
- 3 野中郁次郎、竹内宏高（1996）：93 知識変換モード（SECIモデル）
- 4 ハーズバーグの二要因説はモチベーションに関する理論であり、作業条件や上司などの人間関係は衛生要因とされ、低次の欲求であり、達成感や高い評価などが動機付け要因であり、高次の欲求とされる。高次の欲求が充足されることにより人は動くとする。
- 5 原課の担当者との会話で「忙しくて評価に十分な時

間が割けない」という声が聞かれた。

- 6 原課の評価担当者へのアンケートを2008年3月に実施予定。これにより、特定部署の要因について明らかにできる。

参考文献

- 横山麻季子（2006）「評価による行政サービスの向上」、『日本評価研究』、6（2）、日本評価学会
- 山谷清志（2006）『政策評価の実践とその課題—アカウンタビリティのジレンマ』、萌書房
- 金城慎司（2002）「経済産業省における政策評価の取り組み」、『日本評価研究』、2（2）、日本評価学会
- 上笹恵（2007）「知の融合と経営革新」、日本中小企業学会、同文館
- 野中郁次郎・竹内宏高（1996）『知識創造企業』、東洋経済新報社
- 上野宏「政策工学試論2: 政策プロセス、政策評価、及び予算策定」『日本評価研究』、4（1）、日本評価学会（2008.1.23受理）

Study on the Municipalities Should Do to Improve Evaluation Ability What Municipalities Need to Do

Yasunaga Takachiho

Tamagawa University
takatiho@bus.tamagawa.ac.jp

Abstract

Conducting the right evaluation and providing correct evaluation results to citizens (beneficiaries) are desirable in Municipal policy evaluation. From this point of view, this paper investigates the policy evaluation of Kawasaki City. We found that the departments concluded their program and projects. The committee on the policy evaluation of Kawasaki City, investigated the evaluation report and concluded almost one-third of the reports are difficult to understand. This means the efforts of the implementation departments did not fully achieve its goals.

Kawasaki City should attempt to inform citizens of the evaluation results.

The results of this study shows that municipalities have the motivation and technical skills to do evaluation, but lack the time and support to make effective evaluations. Municipalities also need to develop their evaluation from the standpoint of the citizens.

Keywords

comprehensibility, accountability, organizational correspondent

【研究ノート】

評価の利用における6つのモード

吉澤 剛

サセックス大学科学政策研究所 (SPRU)

gy20@jcom.home.ne.jp

要 約

評価は意思決定や問題解決に資するための戦略的な活動であり、その利用の実態を探ることは学術的にも政策的にも重要な意義を持つ。しかし評価の利用の仕方については様々な見解があり、研究者によって分類や定義がまちまちである。また、日本では理論的・実証的な利用研究がほとんどおこなわれていない。そこで、本稿では利用研究の歴史的背景を振り返り、「利用」という用語について類義語と比較しながら議論した。そして評価の利用という概念に理論的基礎を与える目的で既往文献を広く整理・分析し、利用過程（利用者主導／評価者主導、具体的／抽象的）、利用対象（評価結果／評価の過程）および利用目的（実質的／象徴的）という包括的観点から利用を6つのモードに理論的に分類した。古典的な類型化から離れ、評価の政治化の議論や政策研究からの知見の再評価をおこなうことで、評価の利用における実質性と政治性への平等な視座を獲得した。

キーワード

評価の活用、利用研究、実質の利用、象徴の利用、評価の政治化

1. はじめに

評価の研究・実践はどのように社会の役に立っているのだろうか。たとえば我が国の政策評価について見ると、『政策評価の基本方針』（平成17年12月16日閣議決定）では「政策評価の結果については、各行政機関において、政策評価の結果が政策の企画立案作業（予算要求（定員等を含む。）、法令等による制度の新設・改廃といった作業）における重要な情報として適時的確に活用され、当該政策に適切に反映されるようにする必要がある

」と述べている。これに基づき、どのように政策評価が用いられたかという調査報告が各行政機関¹⁾によってなされている。しかし、こうした評価の反映状況の報告はごく簡単なものにとどまっており、行政機関と独立した学術的な調査研究も乏しい。そもそも日本では評価研究・実践についての反省的な研究がほとんど見られない。評価は意思決定や問題解決に資するための戦略的な活動であり、その政治的文脈に依存する知識を獲得する (Alkin and Taut 2003)。ゆえに他の社会知識や学術研究以上に、評価はその適切な利用を見極め

ねばならない。評価の利用研究に対する社会的要請が近年高まっている理由もそこにある (Alkin 2003)。これは反省的なガバナンス (Voß, Bauknecht and Kemp 2006) という社会科学の新しい流れを汲んでおり、評価研究がより良く社会に貢献するように意識した論究をもっと深めるべきである。

その端緒となるべく、本稿では評価の利用に焦点を当て、利用とは何かについて議論する。まず2章で、評価の利用研究が登場した背景について簡単に振り返る。続く3章では、他の類義語と比較しながら「利用」という用語の概念枠を設定し、利用の分類の目的や視点について明らかにする。そして4章では、既存論文の批判的検討を通じて利用の類型論や評価理論の歴史を追跡しながら、本稿独自の類型を提示し、その意義について考察する。

2. 評価の実用的転回

1950-60年代、欧米政府においては効率的な意思決定やマネジメントのために政策形成や実施ならびにそれらの研究を「科学化」する努力が払われてきた。とりわけ、米国においては福祉や教育サービス、社会プログラムの評価研究に官民間わず巨額が投じられるようになった。しかし、こうした評価は十分に設計されていない単純なものであり、社会状況を考慮したものでもなかった。また当時は、これらの評価の基となる理論的研究も少なかった。評価研究は即時的かつ直接的にプログラムの質の向上に用いられることを意味していたが、1960年代後半になると、評価がプログラムの決定に明確な影響を及ぼすことはないということが経験的に理解され始めた (Rosenblatt 1968; Hofstetter and Alkin 2003, pp.204-205)。そこで一部の研究者は、より実用的な評価のあり方を探求するため、評価活動を政治的な文脈に位置づけることから出発した (Caro 1969; Weiss 1970; cf. Chelimsky 1986)。とりわけWeissは評価現場での実経験を傍証に「評価は政治的文脈で起こる合理的事業である」と言明し、「評価が扱う政策やプログラムは政治的決定の産物である」、「評価は意思決定に資するためにおこなわれるのであるか

ら、その「評価結果」報告は政治的現場に持ち込まれる」、「評価自体が政治的スタンスを取る」ため、評価対象、評価結果の利用、評価自体という三重の意味で評価が政治的であることを示した (Weiss 1973, p.37)。ここで注意しなければならないことは、評価の「実用的転回は、プログラムの中に政治を発見することよりもむしろ「評価」研究は政策形成の実際の過程でより良く用いられるように構築されなければならない」という考えを軸にしている」 (Pawson and Tilley 1997, p.12, 強調は原文)。実際、Weissは評価の政治化についての議論を踏まえて、評価が実質的に役立つ事例について1970年代後半から研究し始めた。そこで彼女が明らかにしたことは、判定や提言として示される評価結果がそのまま意思決定や政策決定に反映されるわけではなく、現実的には意思決定者や政策決定者への「啓蒙 (enlightenment)」あるいは「知識の漸動 (knowledge creep)」という間接的・漸増的な効果に留まることが多いということであった。

1980年代までにWeissが評価の政治性と評価の実質性という二つの側面を提示したことは、評価研究の歴史における重要な里程標となっている。しかし前者が政策研究として発展することを期待され、評価研究の主流から離れて退潮していったことは残念なことである。Weiss自身も政治的な文脈の中でいかに評価を実質化するかについて腐心していたように、両者は本質的に表裏一体である。そこで本稿では、評価の利用研究という架橋的分野を足がかりに、評価研究と政策研究の狭間でおざなりにされてきた評価の政治性に再び光を当ててことを狙いとする。

3. 評価の利用とは何か

評価を実用的なものにするため、1970年代から評価研究の利用についての問題や可能性が探究されるようになり、利用研究 (utilization studies) の発展を担うようになった (Guba 1975; Agarwala-Rogers 1977; Goldstein, Marcus and Rousch 1978)。評価の利用研究は、評価活動がどのように活用されているかについての評価をおこ

なうという意味においてメタ評価として位置づけられる。ニッチな学問分野であるためか、35年余りで100ほどの文献²を数えるのみである。この中で時代を問わず議論の中心となってきたことは、そもそも「利用」とは何かということであり、本稿の主題もそこにある。

まず「利用 (use)」という用語をなぜ用いるのか、他の類義語と比較し、区別したい。利用研究において近年まで最も一般的に用いられてきた「活用 (utilization)」という用語は道具的かつ一時的適用というイメージがある (Weiss 1981, p.18)。評価をこの場で、目に見える形で用いるという意味合いが強く、長期的・間接的に評価が用いられる事例を記述するのに適当でない。日本の評価現場で広く用いられている「反映」という用語も、これと同様の含意がある。「インパクト (impact)」は「大規模なシステムへの間接的な利益、または長期的な便益を示す」(Hofstetter and Alkin 2003, p.216; cf. Bemelmans-Videc 1989-90, p.32) ため、「活用」とは逆の意味で用法が限られる。「効用 (utility)」は主に潜在的な評価の価値を測るときに使われ (Hofstetter and Alkin 2003, pp.216-217)、評価の実際の効果を議論するのに適切な用語とは言えない。さらに「活用」や「効用」という用語は、評価は良いことのためにのみ用いられるという性善的理解を暗に要求するものであり、評価が政治的に利用される文脈には適用しかねる。ここではより中立的な用語である「利用」³を用い、規範的議論には立ち入らない。「影響 (influence)」は「非触知的ないし間接的手段によって他者に対する効果を生み出す人や物の能力」(Kirkhart 2000, p.7) であり、「利用」より広い概念とされる。Cousins & Leithwood (1986, p.332; 1993, p.309) や Weiss (1998, pp.23-24) は「影響」に相当する区分を「利用」の一種類に含めているが、本稿では非意図的で不特定な効果をも内包する「影響」という概念と区別する形で、意図的な効果に概念を限定して「利用」という用語を用いることにする。

利用とは何かを考えると、Patton (1988) が提示したように、いつ、どこで、誰が、何を、どのように、なぜ利用するのかといった視点が挙げられる。Weissが評価の政治的な利用や間接的な

利用に言及したように、ここでは利用の定義に欠かせない「どのように」に論点を絞る。なお、本論では便宜上「誰が」については政策形成や行政の担当者を、「何を」については公共政策評価を想定する。評価がどのように用いられるかを考えることは、評価という理念や行為を再定位することにつながる重要な作業である。本稿では先行研究の批判的レビューを通じて利用にはいろいろな種類があることを示し、類型化する。分類したそれぞれの利用をここではモードと呼ぶことにする。各モードの命名は既存文献に従ったが、概念を新たに定義し直したものもある。Hofstetter & Alkin (2003, pp.204-205) の言うように、利用の概念や操作の狭小さが評価の利用を妨げている側面があるとしたら、より効果的でアカウンタブルな利用のために、利用の古典的概念を拡大し明確な分類をおこなうことが緊要であろう。

ここで注意しなければならないのは、「どのように用いるか」という利用過程は、利用対象（何を）および利用目的（何のために）と密接に関連しているということである。たとえば定規をどのように用いるかは、定規の縁が目盛かという利用対象によって、線を引くため、紙を裁つため、物の長さを測るため、などの利用目的によって変わる。そこで利用の分類にあたり、次章ではまず、評価者主導か利用者主導で用いるか、具体的に用いるか抽象的に用いるか、という利用過程にかかる2つの二分法を導入する。続いて、利用目的として評価を実質的な意図で用いるか象徴的な意図で用いるか、利用対象として評価結果を用いるか評価過程を利用するか、という区分を導入する。最後の節でこれら4つの二分法を一つにまとめた分類表を提示し、議論する。

4. 利用の区分

(1) 評価者主導／利用者主導

まず、評価から意思決定にいたるプロセスが評価者主導か、利用者主導かという区分を考える。これは1980年代後半に勃発したWeiss-Patton論争 (Alkin 1990) の対立構造の基底をなすものと考えら

れる。大まかに両者の主張を追えば、Pattonが利用者に有用な情報を提供することにより合意的で合理的な意思決定を目指したのに対し、Weissは多元的で対立的なコミュニティにおける意思決定は非合理的で政治的なものであり、その文脈で評価者がしなければならないことは正確で十分な評価情報を生み出すことであるとする（Smith and Chircop 1989; Shulha and Cousins 1997）。つまり、Weissが評価者に独立した信頼性を求めたのに対し、Pattonは評価者は利用者の側に立って評価をおこなうべきだと主張している。この区分は認識論的なものだけでなく、Patton（1997）が「実用重視評価」を提唱しているように方法論的にも区分しうる。評価者主導／利用者主導という区分については次節の区分と関係が深いので、次節でも取り上げて議論する。

(2) 具体的／抽象的

具体的に評価結果を利用する（たとえば、評価結果に従ってプログラムの継続の可否を決定する）のか、あるいは抽象的に評価結果を利用する（たとえば、評価結果を何となく参考にして中長期政策を立案する）のかによって利用を分けることができる。具体的利用の中でも「道具的利用」、抽象的利用の中でも「概念的利用」と呼ばれるモードは、評価の利用研究において歴史的に確立された概念となっている。

道具的利用（instrumental use）は意思決定や問題解決のような行動の根拠として、他の手段では得られないような情報を直接引用したり文書化することである（Rich 1977, p.200; 1978, p.101; Beyer and Trice 1982, p.598; Landry, Amara and Lamari 2001, p.336）。この用語を導入した研究者たちは、社会研究は政策問題の解決に資する経験的・結論的証拠を提供する、という工学モデル（engineering model）が道具的利用に包含されるとした（Gouldner 1956; Crawford and Biderman 1969, pp.235-237）。これは、意思決定者は自らの目的についてはっきりした考えや受け入れ可能な代替案を持っており、評価者はそれに必要な情報や知識を《直接》提供できる、という前提に基づく。

「問題解決型」（Havelock 1969, 1971; Rein and White 1977）、「意思決定駆動型」（Weiss 1977; Nelson et al. 1987）、「決定主義者的」（Floden and Weiner 1978）、「離散関数的」（Robertson and Gandy 1983）、「政策駆動型」（Hanney et al. 2003）モデルなどと呼ばれることもある。

しかし、道具の利用は常に工学モデルに従って起こるわけではない。政策形成や行政において道具の利用として理解されるもう一つのモデルは技術官僚モデル（technocratic model）である（Robertson and Gandy 1983; Wittrock 1991）。利用者側が主導する工学モデルと異なり、このモデルでは評価者側が主導的役割を担う（表1）。評価は意思決定のための政治的問題への技術的な解答を促すために設計・利用され（Fischer 1990, p.18）、評価は確定的で直接的な解答を用意できるというのが基本的な前提である。そのため、政策形成者は評価者の専門的な提言に従うのみであり（Foster 1980）、評価結果が自動的に政策に反映されることになる。社会科学でこのモデルに対応するものとしては、「研究・開発・普及（RD&D）」（Havelock 1969, 1971）、「リニア」（Cherns 1972）、「知識駆動型」（Weiss 1977）、「経験主義者的」（Bulmer 1982）、「哲人王」（Coleman 1991）モデルなどがある。

道具の利用の例としては、研究助成の配分機関でおこなわれる事前評価によるプロジェクトの選定が挙げられる。たとえばカナダでは、リサーチカウンシルに設けられた評価委員会によるピアレビューを元に各プロジェクトへの助成金が決定される（Holbrook 2000）。社会知識や社会科学研究的な意思決定における利用を調べたサーベイやインタビューによれば、道具的利用がかなりの割合を占めている（e.g. Caplan, Morrison and Stambaugh 1975; Rich 1977; Knorr 1977; Hutchinson 1995）。こ

表1 利用の4モデル

	評価者主導	利用者主導
道具的利用	技術官僚モデル	工学モデル
概念的利用	啓蒙モデル	古典官僚モデル

（出所）Wittrock (1991, Table 1)を元に作成

れに対し、Weiss (1980) は道具的利用がかなり低い割合にとどまることを示した。この理由としては、利用者を注意深く見極めないかぎり、道具的利用は見つけにくいことが指摘されている (Beyer and Trice 1982, p.599)。

実際はどうであろうか。政治や行政における評価の役割は道具的利用という名の下でなされる活動よりも多面的でとらえがたく、複雑である (Weiss 1979; Wittrock 1991)。現実には、アジェンダや審議、選択といった形をとらないまま評価が利用され、意思決定が徐々に形作られていく (Weiss 1980)。これは**概念的利用** (conceptual use) と呼ばれ、将来の政策形成を目的とする意思決定を、少しずつ間接的に形づくるために、問題についての考え方に影響を及ぼすことを指す (Rich 1977, p.200; 1978, p.101; Weiss 1980, p.382)。社会研究の利用についての調査によると、道具的利用よりも概念的利用が政策形成に大きな影響を与えていることが示唆されている (Caplan, Morrison and Stambaugh 1975; Patton et al. 1977; Rich 1977; Lindblom and Cohen 1979; Weiss 1980; Whiteman 1985a, 1997)。

通常、この概念的利用は啓蒙モデル (enlightenment model) として参照される。啓蒙モデルは工学モデルの対概念とされることもある (Crawford and Biderman 1969, pp.240-241; Janowitz 1970, pp.243-259; Bulmer 1982)。しかし表1に示したように利用の仕方が道具的か概念的か、主導的立場が評価者か利用者か、という二重の意味で対になっていることに注意しなければならない。啓蒙モデルは、評価は概念や提案、方向性、経験的一般化の知的基盤を提供しこそすれ、問題を解決するわけではないという仮定に基づく。どんな単独の評価活動もさしたる効果があるわけではないが、時間とともに利用者に受け入れられるようになる (Weiss 1978)。これは多孔質岩石から滲出する水のように情報が政策変化に間接的・累積的效果をもたらすことから、「石灰岩」モデルとも呼ばれる (Thomas 1985, pp.99-100)。このモデルの一例を挙げよう。オランダでは、大学協会 (VSNU) が大学における研究の質を評価している。その評価は、大学の経営者の意思決定に直接的な影響はないものの、より良い評価を目指す彼

らの考え方や行動に変化を与えているという (Westerheijden 1997)。

利用者主導の概念的利用は古典官僚モデル (classical bureaucratic model) と呼ばれている (Wittrock 1991)。このモデルにおいて、評価は機会や不確実性、競争や他の社会的関係といったコントロールできない社会的プロセスを除外するのではなく、それらと付き合い、共存する。政策立案者はこうした評価に伴う不確定要因を考慮し、意思決定のために間接的に評価を利用するのである。

(3) 実質的／象徴的

上で紹介した道具的利用と概念的利用という区分は、それほどはっきりした形ではないものの1960年代後半までに一部の研究者 (Crawford and Biderman 1969) に認識されていた。1970年代後半に評価の政治化の文脈から象徴的利用という概念が登場すると、道具的・概念的・象徴的利用という三種類の基本的区分が (呼称の違いはあれ) 確立し、評価 (利用) 研究において今日まで参照されている (Pelz 1978; Young and Comtois 1979; Leviton and Hughes 1981; Beyer and Trice 1982; Shulha and Cousins 1997; Landry, Amara and Lamari 2001; Alkin and Taut 2003; Rossi, Lipsey and Freeman 2004, p.411)。象徴的利用 (symbolic use) とは、利用者が自らの見解や既存の立場を正統化し維持するために評価結果を利用することである (Weiss 1984; Pelz 1978; Knorr 1977)。先に議論した評価者主導／利用者主導の区分でいえば、明らかにこれは利用者主導の利用と考えられる。象徴的利用は評価の幅広いインパクトについての不確実性が高く、利用者がその評価を委託しているときに良く見られる。

上述の三区分がこれまで批判的に顧みられてこなかったのは、知識の利用研究あるいは知識政策学と分化するようにして評価研究・評価の利用研究が発展したためと考えられる。これにより、象徴的利用の意味合いについて省察されることがなくなっただと思われる。象徴的利用を道具的・概念的利用の対概念として理論構築する試みは政策学者の Whiteman によってなされたのみであり、こ

こで他の議論を援用しつつ彼の業績を再評価したい。

前節で紹介した道具的利用と概念的利用には、評価の実施者と利用者の間には科学的な評価とそれによる合理的な意思決定という分業が確立されている、という前提がある。これはvan der Knaapらの言う評価における合理的－客観主義的なアプローチと呼応する（van der Knaap 1995; Albæk 1998; Dahler-Larsen 2000）。しかし、評価は純粋な科学的活動ではなく、科学と政治の境界線にあるものである（Weiss 1973）。すなわち利用者は自身の決定や判断に科学的権威を付与しつつも、その決定や判断に対し政治的、行政的、法律的な統制を保っている。これは評価や政策形成に対する主観的－論争的アプローチ（van der Knaap 1995; Albæk 1998; Dahler-Larsen 2000）であり、対立する価値観の中で意思決定が図られる現実を記述している。道具的利用・概念的利用が評価-意思決定プロセスは客観的・合理的なものだという前提に立っているのに対し、象徴的利用はそれを主観的・政治的なものだと見る。したがって、Whitemanらが指摘しているように（Eaton 1962; Whiteman 1982, 1985b）、象徴的利用は道具的利用および概念的利用と別の次元にあると見ることが適切であろう。そこで、彼らにならい、先の二つの利用を実質的利用（substantive use）と呼び、象徴的利用と対置させる。そして道具的利用に対して「説得的利用」、概念的利用に対して「正統化利用」という用語を定義する。

表2に示すように、ここで実質的利用は道具的利用と概念的利用を含み、象徴的利用は説得的利用と正統化利用を含む。

説得的利用（persuasive use）は政論における戦闘手段として評価を利用することである。具体的には評価結果を用いて問題を主張したり、人々に

行動を説得したり、反対者を中立化したり、迷っている人を確信させたり、支持者にテコ入れしたりすることである（Weiss 1978, p.32; Johnson 1998, p.94）。対立を収める手段としてのこうした評価の目的は説得であるため、評価結果を政治的プロパガンダとして活用する傾向にある（Merton and Lerner 1951）。この意味で「政治的」利用と呼ばれることもある（Boeckmann 1976; Weiss 1979）。説得が目的であれば、信頼性の高い証拠を提出するよりも、利用者の好ましい立場を支援するようなやり方で評価がおこなわれてしまい、社会にとって不利益となりうる。証拠があるから説得するのではなく、説得ありきで証拠を揃えるという傾向は、説得的利用のパラドックスと呼ばれている（Henry 2000）。

北欧諸国の研究開発評価についての研究では、以下のように述べて説得的利用の存在を明らかにしている。「ピアレビュー評価が陥りやすいのは、助成を増やすように助言し、評価分野の地位やリソースを高めようとする傾向があることだ」（Luukkonen and Ståle 1990, p.363）。

正統化利用（legitimizing use）とは評価結果を用いて既存の立場を支持あるいは拒否したり、政治的理念や選挙への期待、身びいき、個人的な意向など政治的な考慮により決定や解決を正当化することである（Weiss 1977, p.15; Vedung 1997, p.275; Rossi, Lipsey and Freeman 2004, p.411）。意思決定者が既に持っている意見を代表するような決定を公に支持するために、評価結果は選択的に、また、しばしば歪められて利用される（Knorr 1977）。説得的利用と異なり、評価結果を直接提示して他のアクターに働きかけるというよりは、評価結果を間接的・黙示的に用いて利用者の立場を正当化し、地固めをする（cf. Luukkonen-Gronow 1987）。つまり利用者の立場を積極的に説得するか、それが認められるようそれとなく誘導するかという違いである。

表2 評価結果の利用

	実質的利用	象徴的利用
具体的利用	道具的利用	説得的利用
抽象的利用	概念的利用	正統化利用

（出所）筆者作成

(4) 結果／過程

これまで見てきた利用のモードはすべて結果の利用（findings use）であり、評価は特定のアクターの関心、ニーズ、計算の成果として理解されて

いる (Patton 1997, pp.63-85)。しかしたとえ評価に何の成果がなくても、評価活動の過程が利用されることもある (Patton 1988)。これは過程の利用 (process use) であり、利用対象が異なるために結果の利用と明確に区別できる。評価の利用研究において1980年代後半から登場した概念であり (Shulha and Cousins 1997)、結果と過程の利用を対比させる類型化も一般的となっている (Smith 1988, pp.10-11; Kirkhart 2000, pp.9-10)。結果の利用が総括的評価や形成的評価 (Scriven 1972) としておこなわれるのに対し、過程の利用は構築主義的な評価、すなわち評価とその利用が統合された活動 (Finne, Levin and Nilssen 1995) として「第四世代の評価」 (Guba and Lincoln 1989) と関連づけられる。同じように、Dahler-Larsen (2000) は合理的-客観主義的、主観的-論争的に続く評価のアプローチとして組織の定期的活動に統合される評価のあり方を描いたが、以下に見るように過程の利用には個人の能力向上や、実質的にプロジェクトの質を改善する効果もあり、組織運営のみに貢献するわけではないことを付記しておきたい。

過程の利用という用語を導入し普及させた Patton は、Weiss-Patton 論争の頃から「政治」という言葉を用いることを避けていたように、実質の利用だけを念頭に置いていると見られる。これに対し、本稿では、象徴の利用も含める形でこの用語を広く再定義する。Vedung (1997, pp.274-276) の語用法と Whiteman (1982) の実質的-象徴的区分を適用し、表3に示すように実質的な過程の利用を「対話的利用」、象徴的な過程の利用を「戦術的利用」と呼ぶことにする。

対話的利用 (interactive use) は評価の過程において個人の思考・行動の変化やプログラムや組織の変化を引き起こすことである (Patton 1997, p.90)。個人ないし団体としての利用者が評価にかかわり、評価過程からの経験を解釈、理解し、

意味づけることによって学習をするときにこのモードが出現する。利用者は評価で得た知識をさらなる研究データや常識、一般通念、直感、利用者自身の直接体験などのバックグラウンドと結びつけて利用する。評価への利用者の参加は標準的な評価手続きの明確化と、その改正の両方を促す。この対話的利用は組織的学習、ネットワークの構築、コミュニケーションの拡大、プロジェクトの強化、士気の高揚などの役割を持つ (Forss, Rebien and Carlsson 2002; Patton 1997, 1998)。

過程の利用は、評価の利用に関するサーベイにおいて認められている。そこでは回答者の多くが、内部評価による組織の発展、個人のエンパワメントや自己決定を高く評価している (Preskill and Caracelli 1997)。

意思決定者が「所与の問題に対して《何か》をおこなっている」という姿勢を関係者に示すために評価の過程を利用する場合もある。ここで重要なことは最終的な成果ではなく、評価が任命され、実行されたという事実である。このような利用は**戦術的利用** (tactical use) と呼ばれる。たとえば、なされるべき決定や対策を遅らせたり無視するために評価をおこなうことで時間を稼ぐことがこれにあたる (Knorr 1977, p.171; Vedung 1997, p.276)。また、関係者からの批判をかわすため、不人気な政策成果の責任を回避するため、または政策主体の地位を高めるために評価を利用することも戦術的利用に含まれる (Weiss 1979, 1984; cf. Merton and Lerner 1951)。あるいは客観的で科学的な研究のポーズとして用いられることもある (Suchman 1967, p.143)。いずれにせよ、方法論的困難によるためか、経験的な利用研究では、戦術的利用はほとんど観察されていない (Nilsson 1992)。

(5) まとめ・考察

本章での各分類は表4のようにまとめられる。具体的／抽象的、実質的／象徴的、結果／過程の分類で区分されるものが利用の6つのモードであり、太字で示される。

この類型化が既往研究と異なる特徴は次の通り。(1) 象徴の利用を道具的・概念的利用と別次

表3 評価過程の利用

	実質的利用	象徴的利用
過程の利用	対話的利用	戦術的利用

(出所) 筆者作成

表4 評価の利用の6モード

		実質の利用		象徴の利用
		評価者主導	利用者主導	利用者主導
結果の利用	具体的利用	道具の利用		説得的利用
		技術官僚モデル	工学モデル	
	抽象的利用	概念の利用		正統化利用
		啓蒙モデル	古典官僚モデル	
過程の利用		対話の利用		戦術の利用

(出所) 筆者作成

元の概念であると理論的に位置づけたこと、(2) 説得的・正統化利用という用語を導入し定義づけ、道具的・概念的利用の対概念としたこと、(3) 過程の利用についても実質的／象徴的という区分があるとして、対話の利用と戦術の利用を概念化し、対置させたこと。また、評価者主導／利用者主導という区分を用いれば道具的利用と概念的利用がそれぞれ二つのモデルに分けられることを示した。

なお、実際の文脈では各モードに明確に分類できない場合もあるので注意が必要である。利用過程で考えると、評価者主導／利用者主導については、評価から利用に至る手続きや制度が異なると考えられ、明示的に区分できるだろう。具体的利用と象徴的利用については、これらは排他的ではない。相補的に同時に利用されることもある(Marra 2000)。利用目的は実質的あるいは象徴的な意図を持つものがあるとしたが、実際の利用においては両方の意図が含まれていることが多い(Feldman and March 1981; Pelz 1978)。評価結果の利用と評価過程の利用は利用対象が異なるため、明確に区別できる。このように、評価の利用についての各モードはそれぞれ独立して発現するというのではなく、同時にいくつかのモードが現れる場合もある。

具体的／抽象的、あるいは実質的／象徴的という区分に明確に分類できない場合があると述べたが、これはそれぞれの定義や区分が曖昧であることを意味しない。概念的区分としての厳密性は保たれているが、それらの区分の中間に位置するよ

うな利用や同時発生的な利用が現実的にありうるということである。とはいえ、特に利用が実質的であるか象徴的であるかについては、観察者によって言説が大きく異なる可能性がある。たとえばある評価が政治的に利用されていると研究者が見なしていても、意思決定者がそれを認識していない、あるいは意識的にそうした判断を避けていることは十分に考えられることである。すなわち、評価の利用が政治的でありうるばかりでなく、その利用についての言説も政治的(論争的)になりうる。これからの評価の政治化の議論には、こうした再帰的な見方も含める必要がある。

本稿で扱った評価者主導／利用者主導、具体的／抽象的、実質的／象徴的、結果／過程という二分法はそれぞれ新しい分類法ではないが、これらを組み合わせた類型化は本研究独自の試みである。一部の文献(Weiss 1978, 1979; Vedung 1997; Owen and Rogers 1999, pp.110-113)には明確な区分を示さずに利用の種類を並列しただけのものもあり、これらを整理・統合することがここでの目的でもあった。さらに分類法に混乱が見られる場合もあり、分かりやすい区分を示す必要もあった。たとえば、道具的利用と概念的利用について、Scriven (1996, p.157) は前者が行動指向的で後者が研究指向的と見なせるのではないかと指摘しているが、これは評価者主導／利用者主導という区分と混同したラベリングである。Smith (1988, pp.10-11) は直接的に観察可能／感覚的という具体的／抽象的利用に類似した区分を提唱したが、この分類法は方法論と認識論が混交したものであ

り、理解しにくい。また、Johnson (1998, p.105) は行動的利用を「主に道具的利用だが、象徴的利用、正統化利用、行為指向的過程の利用も含む」とし、それに対して認識的利用を「認識指向的過程の利用、啓蒙的・概念的利用、個人の学習を含む」と定義しているが、雑多な概念が内包されているため非常に分かりにくい。

本稿で取り上げなかった二分法は、即時／長期 (Smith 1988; Kirkhart 2000)、意図的／非意図的 (Kirkhart 2000)、部分的・漸進的／包括的 (Smith 1988)、能動的／受動的 (Westerheijden 1997) である。初めの二つは「影響」と区別するために、「利用」をそれほど長期的なものではなく、また意図的なものに限るためである。加えて、即時的・直接的は具体的、長期的・間接的は抽象的という緩やかな対応関係があり、分類法の重複を避ける必要があった。三番目は利用の仕方というよりも、利用の程度についての議論に強く関係する部分であり、本稿の趣旨から外れるため除外した。最後については、利用者の意識が能動的か受動的かを外部から（また利用者自身も）合理的に認識できるとは考えにくいので、語用法としても分類法としても適当でないと判断した。また、これに相似した評価者主導（利用者にとって受動的側面が強い）／利用者主導（利用者にとって能動的）という区分が既にあることにも配慮した。

過程の利用についての研究はまだ比較的新しく、最近の研究 (Alkin and Taut 2003) で提案しているような道具的（具体的）／概念的（抽象的）という分け方ができるかどうかについて、知見がまだ乏しい。概念的な過程の利用は非意図的であり、評価の利用というよりも影響に相当するのではないかという懸念や、戦術的利用はすべて具体的な利用でしかありえないのではないかという疑問がある。そのためここでは区分をおこなわなかったが、本研究の分類法の課題として残った。今後の実証研究による議論の発展が期待される。

5. 結論

本稿では評価の利用研究という馴染みの薄い分野を取り上げ、評価の利用とは何かについて議論

をおこなった。そして、評価がどのように利用されているかに焦点を当て、既存文献を分析し、利用過程（評価者主導／利用者主導、具体的／抽象的）・利用目的（実質的／象徴的）・利用対象（結果／過程）という包括的な観点からその理論的分類をおこなった。そこで評価の利用を道具的、概念的、説得的、正統化、対話的、戦術的という6つのモードに整理し、それぞれの特徴を明らかにした。この研究の実践的・学術的意義として、まず、評価の評価をおこなうときに、「評価の反映」という即時的・直接的利用の含意が強い概念から離れ、より多角的かつ包括的に評価の意義を捉えることができる。本稿で議論した概念的利用や過程の利用についての理解が深まれば、評価—意思決定の実践がそれほど近視眼的なものではなくなるであろう。また、政府の政策過程における評価の象徴的利用を学術的視点から明らかにすることは、公共政策における評価の位置づけを、より現実 に即したものにする。これまでの評価研究は評価の「客観的」あるいは「科学的」な理論や方法論に固執して、政治的な側面をおざなりにしてきたきらいがあった。本稿では、評価の利用研究で今なお支配的な類型である道具的・概念的・象徴的という三分類を脱し、Weissによる評価の政治化の議論やWhitemanの分類法の再評価をおこなうことで、評価の利用における実質性と政治性への平等な視座を獲得した。これから政治的・社会的文脈における評価の利用についての知見を蓄積することで、評価と利用の構築的關係が深まり、より良い利用を目指した評価研究が活性化するであろう。

本研究でおこなった分類法の今後の展望として、たとえば評価が意思決定・政策過程の「どこで（何に）」「いつ」用いられるかについても議論し、上記の利用の各モードとの関係について考察することが必要であろう。また、今回は行政官などによる公共政策評価の利用を想定したが、他の文脈にどの程度まで一般化できるかについて、理論的・経験的に実証することも求められる。さらに、本稿が評価の利用（アウトプット）について議論したのとは逆に、評価が「なぜ用いられたか」について、利用に影響を与える因子（インプット）を分析するという重要な作業もある。こうした因

子には、問題の本質や評価者・利用者の性質、評価過程や成果の特性、評価者と利用者の対話の質、情報のニーズ、法的・政治的文脈、情報処理のスタイル (Caplan, Morrison and Stambaugh 1975; Patton et al. 1977; Weiss and Bucuvalas 1977, 1980; Sabatier 1978; Johnson 1980; Leviton and Hughes 1981; Beyer and Trice 1982; Alkin 1985; Siegel and Tuckel 1985; Webber 1986, 1987; Cousins and Leithwood 1986, 1993; Nelson et al. 1987; Lester and Wilds 1990; Oh and Rich 1996; Preskill and Caracelli 1997; Owen and Rogers 1999) などが挙げられるが、これについては稿を改めて議論したい。今後はこうしたインプット・アウトプット分析の事例研究を通じて、日本における評価の利用の特徴や評価・意思決定過程・制度について研究をしていきたいと考えている。

注記

- 1 たとえば総務省行政評価局『政策の評価の実施状況及び政策への反映状況』(http://www.soumu.go.jp/s-news/2007/pdf/070608_1_1_H18-0301.pdf) や、内閣府『政策評価結果の政策への反映状況報告書 (平成19年度)』(<http://www8.cao.go.jp/hyouka/h19hanei/houkokusho.pdf>) がある。
- 2 やや古いが、ビブリオグラフィとして、“Utilization of evaluation research” (<http://kuuc.chair.ulaval.ca/english/pdf/bibliographie/evaluations.pdf>) がある。
- 3 日本語では「使用」という用語の方がより中立的かもしれないが、これは対象を直接的に用いるイメージがあり、また「使用 (研究)」より「利用 (研究)」の方が馴染みやすいと考え、ここでは「利用」という用語に統一している。

参考文献

- Agarwala-Rogers, R. (1977). Why is evaluation research not utilized? In M. Guttentag (Ed.), *Evaluation Studies Review Annual*. Beverly Hills, CA: Sage, 327-333.
- Albæk, E. (1998). Knowledge, interests and the many meanings of evaluation: a development perspective. *Scandinavian Journal of Social Welfare*, 7 (2), 94-98.
- Alkin, M. C. (1985). *A Guide for Evaluation Decision Makers*. Beverly Hills, CA: Sage.
- (1990). *Debates on Evaluation*. Newbury Park, CA: Sage.
- (2003). Introduction. In T. Kellaghan & D. L. Stufflebeam (Eds.), *International Handbook of Educational Evaluation*. Dordrecht: Kluwer Academic Publishers, 189-196.
- Alkin, M. C., and Taut, S. M. (2003). Unbundling evaluation use. *Studies in Educational Evaluation*, 29 (1), 1-12.
- Bemelmans-Videc, M.-L. (1989-90). Dutch experience in the utilization of evaluation research: the procedure of reconsideration. *Knowledge in Society: The International Journal of Knowledge Transfer*, 2 (4), 31-48.
- Beyer, J. M., and Trice, H. M. (1982). The utilization process: a conceptual framework and synthesis of empirical findings. *Administrative Science Quarterly*, 27 (4), 591-622.
- Boeckmann, M. E. (1976). Policy impacts of the New Jersey Income Maintenance Experiment. *Policy Sciences*, 7 (1), 53-76.
- Bulmer, M. (1982). *The Use of Social Research: Social Investigation in Public Policy-Making*. London: George Allen & Unwin.
- Caplan, N., Morrison, A., and Stambaugh, R. J. (1975). *The Use of Social Science Knowledge in Policy Decisions at the National Level: A Report to Respondents*. Ann Arbor, MI: Center for Research on Utilization of Scientific Knowledge, Institute for Social Research, University of Michigan.
- Caro, F. G. (1969). Approaches to evaluative research: a review. *Human Organization*, 28 (2), 87-99.
- Chelimsky, E. (1986). What have we learned about the politics of program evaluation? *Evaluation Practice*, 8 (1), 5-21.
- Cherns, A. B. (1972). Models for the use of research. *Human Relations*, 25 (1), 25-33.
- Coleman, D. A. (1991). Policy research - who needs it? *Governance: An International Journal of Policy and Administration*, 4(4), 420-455.
- Cousins, J. B., and Leithwood, K. A. (1986). Current empirical research on evaluation utilization. *Review of Educational Research*, 56 (3), 331-364.

- (1993). Enhancing knowledge utilization as a strategy for school improvement. *Knowledge: Creation, Diffusion, Utilization*, 14 (3), 305-333.
- Crawford, E. T., and Biderman, A. D., eds. (1969). *Social Scientists and International Affairs: A Case for a Sociology of Social Science*. New York: John Wiley & Sons.
- Dahler-Larsen, P. (2000). Surviving the routinization of evaluation: the administrative use of evaluations in Danish municipalities. *Administration & Society*, 32 (1), 70-92.
- Eaton, J. W. (1962). Symbolic and substantive evaluative research. *Administrative Science Quarterly*, 6 (4), 421-442.
- Feldman, M. S., and March, J. G. (1981). Information in organizations as signal and symbol. *Administrative Science Quarterly*, 26 (2), 171-186.
- Finne, H., Levin, M., and Nilssen, T. (1995). Trailing research: a model for useful program evaluation. *Evaluation*, 1 (1), 11-31.
- Fischer, F. (1990). *Technocracy and the Politics of Expertise*. Newbury Park, CA: Sage.
- Floden, R. E., and Weiner, S. S. (1978). Rationality to ritual: the multiple roles of evaluation in governmental processes. *Policy Sciences*, 9(1), 9-18.
- Forss, K., Rebien, C. C., and Carlsson, J. (2002). Process use of evaluations: types of use that precede lessons learned and feedback. *Evaluation*, 8(1), 29-45.
- Foster, J. L. (1980). An advocate role model for policy analysis. *Policy Studies Journal*, 8 (6), 958-964.
- Goldstein, M. S., Marcus, A. C., and Rousch, N. P. (1978). The nonutilization of evaluation research. *Pacific Sociological Review*, 21 (1), 21-44.
- Gouldner, A. W. (1956). Explorations in applied social science. *Social Problems*, 3(3), 169-181.
- Guba, E. G. (1975). Problems in utilizing results of evaluation. *Journal of Research and Development in Education*, 8(3), 42-54.
- Guba, E. G., and Lincoln, Y. S. (1989). *Fourth Generation Evaluation*. Newbury Park: Sage.
- Hanney, S. R., Gonzalez-Block, M. A., Buxton, M. J., and Kogan, M. (2003). The utilisation of health research in policy-making: concepts, examples and methods of assessment. *Health Research Policy and Systems*, 1(2). Available at <http://www.health-policy-systems.com/content/1/1/2>.
- Havelock, R. G. (1969). *Planning for Innovation through Dissemination and Utilization of Knowledge*. Ann Arbor, MI: Center for Research on Utilization of Scientific Knowledge, Institute for Social Research, The University of Michigan.
- (1971). The utilisation of educational research and development. *British Journal of Educational Technology*, 2 (2), 84-98.
- Henry, G. T. (2000). Why not use? In V. J. Caracelli & H. Preskill (Eds.), *The Expanding Scope of Evaluation Use*. San Francisco, CA: Jossey-Bass, 85-98.
- Hofstetter, C. H., and Alkin, M. C. (2003). Evaluation use revisited. In T. Kellaghan & D. L. Stufflebeam (Eds.), *International Handbook of Educational Evaluation*. Dordrecht: Kluwer Academic Publishers, 197-222.
- Holbrook, A. (2000). Evaluation of research sponsored by federal granting councils in Canada: the social contract. *Research Evaluation*, 8 (1), 47-56.
- Hutchinson, J. R. (1995). A multimethod analysis of knowledge use in social policy: research use in decisions affecting the welfare of children. *Science Communication*, 17 (1), 90-106.
- Janowitz, M. (1970). *Political Conflict: Essays in Political Sociology*. Chicago: Quadrangle Books.
- Johnson, K. W. (1980). Stimulating evaluation use by integrating academia and practice. *Knowledge: Creation, Diffusion, Utilization*, 2 (2), 237-262.
- Johnson, R. B. (1998). Toward a theoretical model of evaluation utilization. *Evaluation and Program Planning*, 21 (1), 93-110.
- Kirkhart, K. E. (2000). Reconceptualizing evaluation use: an integrated theory of influence. In V. J. Caracelli & H. Preskill (Eds.), *The Expanding Scope of Evaluation Use*. San Francisco, CA: Jossey-Bass, 5-23.
- Knorr, K. D. (1977). Policymakers' use of social science knowledge: symbolic or instrumental? In C. H. Weiss (Ed.), *Using Social Research in Public Policy Making*. Lexington, MA: Lexington Books, 165-182.
- Landry, R., Amara, N., and Lamari, M. (2001). Utilization of social science research knowledge in Canada. *Research*

- Policy*, 30(2), 333-349.
- Lester, J. P., and Wilds, L. J. (1990). The utilization of public policy analysis: a conceptual framework. *Evaluation and Program Planning*, 13(3), 313-319.
- Leviton, L. C., and Hughes, E. F. X. (1981). Research on the utilization of evaluations: a review and synthesis. *Evaluation Review*, 5(4), 525-548.
- Lindblom, C. E., and Cohen, D. K. (1979). *Usable Knowledge: Social Science and Social Problem Solving*. New Haven and London: Yale University Press.
- Luukkonen, T., and Stähle, B. (1990). Quality evaluations in the management of basic and applied research. *Research Policy*, 19(4), 357-368.
- Luukkonen-Gronow, T. (1987). Scientific research evaluation: a review of methods and various contexts of their application. *R&D Management*, 17(3), 207-221.
- Marra, M. (2000). How much does evaluation matter?: some examples of the utilization of the evaluation of the World Bank's anti-corruption activities. *Evaluation*, 6(1), 22-36.
- Merton, R. K., and Lerner, D. (1951). Social scientists and research policy. In D. Lerner & H. D. Lasswell (Eds.), *The Policy Sciences*. Palo Alto, CA: Stanford University Press, 282-307.
- Nelson, C. E., Roberts, J., Maederer, C. M., Wertheimer, B. and Johnson, B. (1987). The utilization of social science information by policymakers. *American Behavioral Scientist*, 30(6), 569-577.
- Nilsson, K. (1992). Utilization patterns and strategies in three policy sectors. *Science Studies*, 5(1), 29-46.
- Oh, C. H., and Rich, R. F. (1996). Explaining use of information in public policymaking. *Knowledge and Policy: The International Journal of Knowledge Transfer and Utilization*, 9(1), 3-35.
- Owen, J. M., and Rogers, P. J. (1999). From evaluation findings to utilisation, *Program Evaluation: Forms and Approaches*. Sage, 105-131.
- Patton, M. Q. (1988). Six honest serving men for evaluation. *Studies in Educational Evaluation*, 14(3), 301-330.
- (1997). *Utilization-Focused Evaluation: The New Century Text 3rd edition*. Thousand Oaks, CA: Sage.
- (1998). Discovering process use. *Evaluation*, 4(2), 225-233.
- Patton, M. Q., Grimes, P. S., Guthrie, K. M., et al. (1977). In search of impact: an analysis of the utilization of federal health evaluation research. In C. H. Weiss (Ed.), *Using Social Research in Public Policy Making*. Lexington Books, 141-163.
- Pawson, R., and Tilley, N. (1997). *Realistic Evaluation*. London: Sage.
- Pelz, D. C. (1978). Some expanded perspectives on use of social science in public policy. In J. M. Yinger & S. J. Cutler (Eds.), *Major Social Issues: A Multidisciplinary View*. New York: Free Press, 346-357.
- Preskill, H., and Caracelli, V. J. (1997). Current and developing conceptions of use: evaluation use TIG survey results. *Evaluation Practice*, 18(3), 209-225.
- Rein, M., and White, S. H. (1977). Can policy research help policy? *Public Interest*, 49, 119-136.
- Rich, R. F. (1977). Uses of social science information by federal bureaucrats: knowledge for action versus knowledge for understanding. In C. H. Weiss (Ed.), *Using Social Research in Public Policy Making*. Lexington, MA: Lexington Books, 199-211.
- (1978). The use of science in policymaking: a comparative perspective on science policy. In D. E. Ashford (Ed.), *Comparing Public Policies: New Concepts and Methods*. Beverly Hills and London: Sage, 99-110.
- Robertson, A., and Gandy, J. (1983). Policy, practice and research: an overview. In J. Gandy, A. Robertson & S. Sinclair (Eds.), *Improving Social Intervention: Changing Social Policy and Social Work Practice through Research*. London and Canberra: Croom Helm, 241-289.
- Rosenblatt, A. (1968). The practitioner's use and evaluation of research. *Social Work*, 13(1), 53-59.
- Rossi, P. H., Lipsey, M. W., and Freeman, H. E. (2004). *Evaluation: A Systematic Approach 7th edition*. Thousand Oaks, CA: Sage.
- Sabatier, P. (1978). The acquisition and utilization of technical information by administrative agencies. *Administrative Science Quarterly*, 23(3), 396-417.
- Scriven, M. (1972). The methodology of evaluation. In C. H. Weiss (Ed.), *Evaluating Action Programs: Reading in Social Action and Education*. Boston, MA: Allyn and Bacon, 123-136.
- (1996). Types of evaluation and types of evaluator. *Evaluation Practice*, 17(2), 151-161.

- Shulha, L. M., and Cousins, B. (1997). Evaluation use: theory, research, and practice since 1986. *Evaluation Practice*, 18(3), 195-208.
- Siegel, K., and Tuckel, P. (1985). The utilization of evaluation research: a case analysis. *Evaluation Review*, 9(3), 307-328.
- Smith, M. F. (1988). Evaluation utilization revisited. In J. A. McLaughlin, L. J. Weber, R. W. Covert & R. B. Ingle (Eds.), *Evaluation Utilization*. San Francisco, CA: Jossey-Bass, 7-19.
- Smith, N. L., and Chircop, S. (1989). The Weiss-Patton debate: illumination of the fundamental concerns. *American Journal of Evaluation*, 10(1), 5-13.
- Suchman, E. A. (1967). *Evaluative Research: Principles and Practice in Public Service & Social Action Programs*. New York: Russell Sage Foundation.
- Thomas, P. (1985). *The Aims and Outcomes of Social Policy Research*. London: Croom Helm.
- Van der Knaap, P. (1995). Policy evaluation and learning: feedback, enlightenment or argumentation? *Evaluation*, 1(2), 189-216.
- Vedung, E. (1997). *Public Policy and Program Evaluation*. New Brunswick and London: Transaction Publishers.
- Voß, J.-P., Bauknecht, D., and Kemp, R., eds. (2006). *Reflexive Governance for Sustainable Development*. Cheltenham: Edward Elgar.
- Webber, D. J. (1986). Explaining policymakers' use of policy information: the relative importance of the two-community theory versus decision-maker orientation. *Knowledge: Creation, Diffusion, Utilization*, 7(3), 249-290.
- (1987). Legislators' use of policy information. *American Behavioral Scientist*, 30(6), 612-631.
- Weiss, C. H. (1970). The politicization of evaluation research. *Journal of Social Issues*, 26(4), 57-68.
- (1973). Where politics and evaluation research meet. *Evaluation*, 1(3), 37-45.
- (1977). Introduction. In C. H. Weiss (Ed.), *Using Social Research in Public Policy Making*. Lexington Books, 1-22.
- (1978). Improving the linkage between social research and public policy. In L. E. Lynn, Jr. (Ed.), *Knowledge and Policy: The Uncertain Connection*. Washington, DC: National Academy of Sciences, 23-81.
- (1979). The many meanings of research utilization. *Public Administration Review*, 39(5), 426-431.
- (1980). Knowledge creep and decision accretion. *Knowledge: Creation, Diffusion, Utilization*, 1(3), 381-404.
- (1981). Measuring the use of evaluation. In J. A. Ciarlo (Ed.), *Utilizing Evaluation: Concepts and Measurement Techniques*. Beverly Hills, CA: Sage, 17-33.
- (1984). Increasing the likelihood of influencing decisions. In L. Ruttman (Ed.), *Evaluation Research Methods: A Basic Guide*. Beverly Hills, CA: Sage, 159-190.
- (1998). Have we learned anything new about the use of evaluation? *American Journal of Evaluation*, 19(1), 21-33.
- Weiss, C. H., and Bucuvalas, M. J. (1977). The challenge of social research to decision making. In C. H. Weiss (Ed.), *Using Social Research in Public Policy Making*. Lexington Books, 213-233.
- (1980). Truth tests and utility tests: decision-makers' frames of reference for social science research. *American Sociological Review*, 45(2), 302-313.
- Westerheijden, D. F. (1997). A solid base for decisions. *Higher Education*, 33(4), 397-413.
- Whiteman, D. (1982). Congressional use of technology assessment. In D. M. O'Brien & D. A. Marchand (Eds.), *The Politics of Technology Assessment: Institutions, Processes, and Policy Disputes*. Lexington, MA: Lexington Books, 51-64.
- (1985a). The fate of policy analysis in congressional decision making: three types of use in committees. *Western Political Quarterly*, 38(2), 294-311.
- (1985b). Reaffirming the importance of strategic use. *Knowledge: Creation, Diffusion, Utilization*, 6(3), 203-224.
- (1997). Congress and policy analysis: a context for assessing the use of OTA projects. *Technological Forecasting and Social Change*, 54(2/3), 177-189.
- Wittrock, B. (1991). Social knowledge and public policy: eight models of interaction. In P. Wagner, C. H. Weiss, B. Wittrock & H. Wollmann (Eds.), *Social Sciences and Modern States: National Experiences and Theoretical*

Crossroads. Cambridge: Cambridge University Press, 333-353.

Young, C. J., and Comtois, J. (1979). Increasing congressional utilization of evaluation. In F. M. Zweig

(Ed.), *Evaluation in Legislation*. Beverly Hills & London: Sage.

(2008.1.24受理)

Six Modes of Evaluation Use

Go Yoshizawa

SPRU - Science and Technology Policy Research, University of Sussex
gy20@jcom.home.ne.jp

Abstract

A vast reservoir of knowledge and experience from utilization studies developed since 1970s illuminates political and enlightenment functions in evaluation. As a bridgehead for applying the concept of evaluation use to evaluation studies and practices for public policy in Japan, this paper starts with an historical review of utilization studies and frames the concept of ‘use’ distinguishing the term from its synonyms. Refining a number of definitions and perspectives from earlier studies, it then introduces evaluator-primacy/user-primacy, concrete/abstract (process of use), substantive/symbolic (purpose of use) and findings/process (object of use) discriminators. With these discriminators and the appreciation of studies by policy analysts such as Weiss and Whiteman, this study illustrates six different modes of evaluation use with a novel comprehensive matrix. Apart from the conventional typology of use including instrumental, conceptual and symbolic, this matrix gains a balanced perspective on the substantive and political nature of evaluation use.

Keywords

evaluation utilization, utilization studies, substantive use,
symbolic use, politicization of evaluation

日本評価学会第8回全国大会 開催報告

1. 第8回全国大会プログラム（実績）

2. 共通論題セッション報告

共通論題 1 「評価のバウンダリー」	座長	三好 皓一（立命館アジア太平洋大学）
共通論題 2 「キャパシティ・ディベロップメントの評価」	座長	和田 義郎（政策研究大学院大学）
	討論者	新海 尚子（名古屋大学大学院）
共通論題 3 「評価教育論」	座長	岩渕 公二（参議院）
	討論者	西山 慶司（法政大学大学院特定課題研究所）
共通論題 4 「自治体評価のキャパシティ・ディベロップメント」	座長	高千穂安長（玉川大学）
	討論者	山谷 清志（同志社大学）
共通論題 5 「介護サービスの評価」	座長	大島 巖（日本社会事業大学）
	座長	平岡 公一（お茶の水女子大学）
共通論題 6 「『地域おこし』の評価」	座長	林 薫（文教大学）
	討論者	梅田 次郎（日本能率協会コンサルティング）

3. シンポジウム報告

シンポジウム「中期目標・中期計画の評価」	司会	松本 哲男（名古屋大学）
----------------------	----	--------------

4. 自由論題セッション報告

自由論題 1 「市民社会」	座長	入山 映（サイバー大学）
自由論題 2 「政策評価」	座長	小野 達也（鳥取大学）
自由論題 3 「教育評価」	座長	石田 洋子（国際開発センター）
自由論題 4 「行政評価」	座長	宮崎 修二（対日貿易投資交流促進協会）
自由論題 5 「国際協力 1」	座長	牟田 博光（東京工業大学）
自由論題 6 「国際協力 2」	座長	林 薫（文教大学）

日本評価学会第8回全国大会 開催のご報告とお礼

2007年11月10-11日開催（於：名古屋大学）の日本評価学会第8回全国大会には110名を超える多くの方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お願い申し上げます。

第8回全国大会

実行委員会委員長

プログラム委員会委員長

松本 哲男

牟田 博光

日本評価学会第8回全国大会
「評価結果をどう活かすか」(於：名古屋大学)

第1日：2007年11月10日（土）

9：15-9：45	受 付		
午前の部 9：45-11：45	共通論題 1 評価のバウンダリー (三好皓一)	自由論題 1 市民社会 (入山映)	
お昼休み 11：45-12：45	お昼休み（理事会 11：45 - 12：45）		
午後の部Ⅰ 12：45-14：45	共通論題 2 キャパシティ・ ディベロップメントの評価 (和田義郎)	自由論題 2 政策評価 (小野達也)	自由論題 3 教育評価 (石田洋子)
午後の部Ⅱ 15：00-17：00	シンポジウム 「中期目標・中期計画の評価」 (松本哲男)		
17：10-17：50	総 会		
18：00-20：00	懇親会		

第2日：2007年11月11日（日）

9：30-10：00	受 付		
午前の部 10：00-12：00	共通論題 3 評価教育論 (岩渕公二)	自由論題 4 行政評価 (宮崎修二)	自由論題 5 国際協力1 (牟田博光)
お昼休み 12：00-13：15	お昼休み		
午後の部 I 13：15-15：15	共通論題 4 自治体評価のキャパシティ・ ディベロップメント (高千穂安長)	共通論題 5 介護サービスの評価 (大島巖)	自由論題 6 国際協力2 (林薫)
午後の部 II 15：30-17：00	共通論題 6 「地域おこし」の評価 (林薫)		特別セッション 評価インターンによる中間報告 (西出順郎)

（ ）内は座長

自由論題報告の時間は原則として、発表20分、質疑8分、交替2分（計30分）です。

日本評価学会第8回全国大会 「評価をどう活かすか」プログラム

11月10日（土）午前の部 9：45-11：45

共通論題 1 評価のバウンダリー			
		座長	三好 皓一 立命館アジア太平洋大学
評価研究は専門分野（Discipline）として確立しうるのか？ ー学問としての評価研究を考えるー		松岡 俊二	早稲田大学
日本型の評価とは何か ー日本型政策評価についてー		窪田 好男	神戸学院大学
問われる評価体系と政策マネジメント・サイクル		田中 弥生	大学評価・学位授与機構
評価制度、評価体制はどのように扱うべきか？ ー評価論として取り込んでいく重要性についてー		三好 皓一	立命館アジア太平洋大学
自由論題 1 市民社会			
		座長	入山 映 サイバー大学
新傾向のCSR報告書の評価に関する考察		井上 昌美	筑波大学大学院
The Japanese Red Cross Society's Organization; its Independence or Conformity with Government Structure		Yui, Masashi	Meiji University / Japanese Red Cross Society
市民社会による日本の対アフリカ政策評価のあり方に関する検討（II）		石田 洋子	TICAD市民社会フォーラム

11月10日（土）午後の部 I 12：45-14：45

共通論題2 キャパシティ・ディベロップメントの評価			
	座長 討論者	和田 義郎 新海 尚子	政策研究大学院大学 名古屋大学大学院
キャパシティ・ディベロップメント（CD）評価の可能性を考える ーCDアプローチの展開と社会的能力アセスメント（SCA）ー		松岡 俊二	早稲田大学
CDの視点と技術協力プロジェクトの評価		三浦 和紀	国際協力機構
円借款事業におけるCDとその評価	○	川淵 貴代	国際協力銀行
		下村 暢子	コーエイ総合研究所
自由論題2 政策評価			
	座長	小野 達也	鳥取大学
自治体政策評価の活用インセンティブ体系と個人レベルの活用に関する考察		福嶋 隆宏	高崎経済大学大学院
政策評価における中央地方関係ー政府間評価の試みー		西本 哲也	同志社大学大学院
冬期道路管理の業績評価とマネジメントについて		浅野 基樹	土木研究所 寒地土木研究所
市場化テストのPDCAに見る評価手法の考察ーマニフェスト評価体系からのアプローチー	○	田中 弥生	大学評価・学位授与機構
		工藤 泰志	（特）言論NPO
自由論題3 教育評価			
	座長	石田 洋子	国際開発センター
中学校における授業評価の研究ー教員の授業評価に対する意識に着目してー	○	市野 緑	東京工業大学大学院
		牟田 博光	東京工業大学大学院
高等教育におけるe-Learningの効果に関する研究結果のメタ分析	○	金 性希	東京工業大学大学院
		齋藤 貴浩	大学評価学位授与機構
TIMSS2003における理科の学力に関する要因分析ーヨルダンを事例としてー	○	藤谷 元子	東京工業大学大学院
		牟田 博光	東京工業大学大学院
大学生生活満足度の要因重要度に関する分析		星野 敦子	十文字学園女子大学

11月10日（土）午後の部Ⅱ 15:00-17:00

シンポジウム 中期目標・中期計画の評価		座長 松本 哲男 名古屋大学	
	パネリスト	藤原 誠	文部科学省高等教育局高等教育企画課長
	パネリスト	伊藤 直樹	外務省国際協力局政策課課長
	パネリスト	栗本 英和	名古屋大学教授・評価企画室副室長
	パネリスト	山谷 清志	同志社大学教授・総務省独立行政評価委員会委員・日本評価学会副会長
	討論者	牟田 博光	東京工業大学理事/副学長・日本評価学会副会長

11月10日（土）総会 17:10-17:50

11月10日（土）懇親会 18:00-20:00

11月11日（日）午前の部 10:00-12:00

共通論題 3 評価教育論		座長 岩淵 公二 参議院 討論者 西山 慶司 法政大学大学院特定課題研究所	
わが国における政策評価の研究と教育		山谷 清志	同志社大学
評価教育と公共政策学		窪田 好男	神戸学院大学
政策系学部における評価教育のあり方 ー技術の学か、社会の技術かー		南島 和久	長崎県立大学
自由論題 4 行政評価		座長 宮崎 修二 （財）対日貿易投資交流促進協会	
経営の進捗発信による学校の信頼度向上に関する実践報告		石田 謙豪	広島県尾道市立栗原北小学校
行政評価制度のパフォーマンスの概念とその分析手法に関する考察		田中 啓	静岡文化芸術大学
米国連邦政府におけるProgram Assessment and Rating Tool: PART運用 状況について～業績と予算とのリンケージに関する考察～		左近 靖博	三菱UFJリサーチ&コンサルティング
自由論題 5 国際協力 1		座長 牟田 博光 東京工業大学	
Metaevaluation of Aid Evaluations		Sasaki, Ryo	International Development Center of Japan / Western Michigan University
「目標から自由な評価」考察		藤田 伸子	国際開発高等教育機構
参加型評価による援助パラドックスの解決		田中 博	ヒマラヤ保全協会
大学による国際協力機関との連携について ー事業受託の問題点とその解決に向けてー	○	杉本 充邦	名古屋大学
		松本 哲男	名古屋大学

11月11日（日） 午後の部Ⅰ 13:15-15:15

共通論題 4 自治体評価のキャパシティ・ディベロップメント

座長 高千穂安長 玉川大学
討論者 山谷 清志 同志社大学

地域の経営と評価能力		岩渕 公二	参議院
自治体評価における計量・計数の現状と課題		小野 達也	鳥取大学
地域社会における評価のキャパシティディベロップメント		鎌田 徳幸	岩手県庁
自治体の評価能力キャパシティビルディング		高千穂安長	玉川大学

共通論題 5 介護サービスの評価

座長 大島 巖 日本社会事業大学
座長 平岡 公一 お茶の水女子大学

認定データを用いた介護サービスの評価の考え方と実際		筒井 孝子	国立保健医療科学院
高齢者・家族からみた介護保険制度のアウトカム評価		杉澤 秀博	桜美林大学
介護サービスの自立支援効果	○	菊澤佐江子	奈良女子大学
		澤井 勝	奈良女子大学

自由論題 6 国際協力 2

座長 林 薫 文教大学

ベトナムの中小企業支援プロジェクトに関する評価	○	菊池 正 大平 哲	国際開発高等教育機構 慶應義塾大学
ネパール政府による公共事業の評価システム強化と日本の支援のあり方について	○	石田 洋子 三浦 順子	国際開発センター グローバル・リンク・マネジメント
Localization of Evaluation — Experiences from Nepal and the Philippines	○	Wachowski Cindy Lyn Miyoshi, Koichi	Ritsumeikan Asia Pacific University Ritsumeikan Asia Pacific University

11月11日（日） 午後の部Ⅱ 15:30-17:30

共通論題 6 「地域おこし」の評価

座長 林 薫 文教大学
討論者 梅田 次郎 日本能率協会コンサルティング

Evaluating Community Capacity Development: The Case of Oyama-cho	○	Stenning, Naomi Miyoshi, Koichi	Ritsumeikan Asia Pacific University Ritsumeikan Asia Pacific University
地域おこしの評価方法—非市場財を用いた評価方法—	○	田名邊賢治	総務省行政評価局
	○	小林 優輔	総務省行政評価局

特別セッション 評価インターンによる調査報告（中間報告）

座長 西出順郎 琉球大学

JICAシリア国リプロダクティブヘルス強化プロジェクトの評価	○	松本 知子	早稲田大学大学院
	○	村田 維沙	東京工業大学大学院
<派遣先：（株）アースアンドヒューマンコーポレーション>	討論者	源 由理子	明治大学
倉敷青年会議所の事業評価と評価マニュアルの作成	○	高田 岳志	立教大学大学院
<派遣先：（社）倉敷青年会議所>	○	橋本 裕人	早稲田大学大学院
	討論者	高千穂安長	玉川大学

共通論題セッション報告

共通論題セッション1「評価のバウンダリー」

座長 三好 皓一（立命館アジア太平洋大学）

日本評価学会は、評価学会設立趣意書において、「本学会の事業内容は、対象分野における研究成果の実践性を重視して行われるものとする。」とし「研究分野としては、・評価の体系（評価ニーズとそれに対応する評価手法の整理・体系化に関する研究）、・評価の手法（評価手法の高度化のための研究）、・評価の技術（評価実施の技術的側面に関する研究）に重点を置くこととする。」と記載し設立された。このような研究分野を設定し設立された日本評価学会は、設立以来約7年を経て、評価学会員数は現在495名（2007/8/1）となり、また、全国大会、評価学会誌等によって多くの評価について議論が行ってきた。このような状況をふまえ、7年を経たこの時期に1つの区切りとして、評価についての議論をどのような範囲で行っていくのか、すなわち評価のバウンダリーについて議論し、研究対象である評価論について少しでも明確にすることを目的に本評価セッションは構成された。

座長の三好より、趣旨説明および問題提起がなされ、報告者4名が報告を行なった。松岡報告（早稲田大学）：「評価研究は専門分野（Discipline）として確立しうるのか？：学問としての評価研究を考える」は、評価研究の研究分野としての位置づけ、学問的な特性、日本社会の中での評価研究の専門学問分野（academic discipline）としての確立可能性について考察した。欧米を中心とする世界の評価研究の動向を概観し、評価研究の制度化の進展と評価研究における科学性と実用性の流れを明らかにするとともに、日本における評価研究の流れを振り返り、日本における新たな学際的・複合的な学問体系創設の困難性と可能性という観点から評価研究の今後のあり方を提示した。窪田報告（神戸学院大学）：「日本型政策評価とは何か」は、日本型政策評価を手掛かりとして日本社会の文脈において評価論を明確にするにあたっての論点・視点を提示した。本報告は、地方自治体の事務事業評価について、日本型政策評価としての、また、体系的評価を目指した職員による自己評価としての特質とともにその実行可能性を指摘し、職員による自己評価でこそ正確に明らかにできる事項に特化し、公共政策のデザインの活性化にその目的を変更することを求めた。田中報告（大学評価・学位授与機構）：「問われる評価体系と政策マネジメント・サイクル」は、日本の政策評価の限界について問題提起しその解決の方向性を探った。現行の政策評価の課題として、・政治レベルの判断で策定された複数の府省に横たわる政策の一省庁による政策評価の限界性、・府省の上位にある政治レベルの判断や影響による政策評価結果の無視と政策マネジメント・サイクルの崩壊、・アカウンタビリティの対象の曖昧性と欠如の3点を挙げ、評価における政策評価体系と政策マネジメント・サイクルのあり方について考察した。最後に、三好報告：「評価制度、評価体制はどのように扱うべきか？：評価論として取り込んでいく重要性について」は、ガバメントからガバナンスへの移行に焦点を当て、安定的な構造を前提とした従来の方法論を主体とした評価の議論を超えた、流動的な構造の下で実社会における評価のあり方について議論した。特に、評価体制の評価論としての取り扱いとその課題について指摘した。

会場からは多くの質問とコメントがあった。特に、学会員構成が研究者に偏ってきており実務家の影が薄くなっていることの意味、また、評価結果を予算編成に結び付けていく道筋についてなど学会のあり方、評価の実践性の視点から議論を行なった。

これらの報告と議論は、我が国の評価研究の対象についての示唆を提示するものであり、今後の評価論

の議論を展開する上で有意義なものであった。

共通論題セッション2「キャパシティ・ディベロップメントの評価」

座長 和田 義郎（政策研究大学院大学）

討論者 新海 尚子（名古屋大学大学院）

本セッションは以下の3件の発表の後、質疑応答があった。

「キャパシティ・ディベロップメント（CD）評価の可能性を考える－CDアプローチの展開と社会的能力アセスメント－」（松岡俊二）では、CD評価の成立しうる手法・方法論とその評価可能性をCD概念まで遡って検討し、UNDPが置換アプローチの限界から個人、組織、社会が相互に関連しながら能力が形成されるとするCDを提案したこと、むしろ、政府、企業、市民の3者を社会的アクターとした社会的能力（SCD）に着眼すべきであること、その上で、この社会的能力をアクターファクター分析や指標化、制度分析、経路分析、発展ステージ分析などの手法で評価してゆくことが必要となること、そしてその適用事例としてインドネシアとモンゴルの例が報告された。

「CDの視点と技術協力プロジェクトの評価」（三浦和紀）では、JICAにおいては、途上国の課題対処能力が、個人、組織、社会などの複数のレベルの総体として向上してゆくことをCDとしている点を踏まえ、内発性、包括性の二つの観点から、JICA技術協力プロジェクトを5つの目標の類型から分類し、自立発展性やプログラムの思考の必要性との関係で、どのように評価すべきか、が報告された。

「円借款事業におけるCDとその評価」（川淵貴代、下村暢子）では、円借款のCDについて、とくにJICA連携の評価において5つの類型が存在すること、そして、それぞれの中のケーススタディ（ベトナム、ヨルダン、インドネシア等）が報告され、それぞれの中での困難や問題点、逆に良い点などをまとめた上での円借款事業におけるCDの評価課題が報告された。

これらの発表に対し、討議者（新海尚子）より、キャパシティの概念やCDの評価可能性といった論点について議論が提起され、その後、会場から、数量的に計測する場合の実務の枠組みはどのようなものか、特に能力の開発の度合いをどのように計測するのか、実際には、事前調査の段階で組み込むべきではないか、内発性とカントリー・システムとくに事業実施単位との関係、CDは失敗事例についてより顕著に明確となるのではないかなどについて質疑があった。

キャパシティ・ディベロップメントの評価は、社会的能力の形成を評価し、それを援助の効果として捉えるという意味で肯定的に考えられる一方で、とくにCD評価の手法、実際上の計測可能性などにより、現在の段階での課題も存在しており、キャパシティ・ディベロップメントの評価の可能性と今後の課題を十分に検討してゆくことが現実の実務への適用には必要であることを改めて感じさせられた。

共通論題セッション3「評価教育論」

座長 岩淵 公二（参議院）

討論者 西山 慶司（法政大学大学院特定課題研究所）

本セッションの目的は、大学における評価教育のデザインと実践などを踏まえ、政策学における政策評価論の教育のあり方およびその問題点を明らかにすることであった。

わが国では、ここ10年の間に法学部、経済学部を中心に公共政策論を科目として提供する大学が増えてきた。しかしながら、いまだ十分に政策評価論の科目が普及したとはいえない。そこで、報告者には、大学のとくに学部段階を念頭に置きながらカリキュラム、教科書、教育人材、学問ディシプリン等の視点か

ら発表いただき、政策学における評価教育の実態と課題について議論を進めた。

第一報告者の山谷清志会員（同志社大学政策学部）は、政策評価の教科書の不在に政策評価エキスパートを育成する困難の一端があることを指摘したうえで、政策評価の思考をする手順として「政策の対象」「政策学」「制度と管理」「スキル」の4つの項目について具体的に示し、政策学部カリキュラムにおける政策評価論講義の構想を披露した。第二報告者の南島和久会員（長崎県立大学経済学部）は、政策評価論の講義が普及していない実態に触れ、その原因を「政策評価の実務偏重の状況」「学問間の役割分担の状況」「公共政策教育や評価教育の受けとめられ方」「教科書等の体制整備の状況」「社会のニーズや評価教育をとりまく環境等」にあるとしたうえで、政策評価教育の確立に向けた継続的な意見交換の場の必要性を提起した。第三報告者の窪田好男会員（神戸学院大学法学部）は、既存のディシプリンの手法をマルチ・ディシプリンで活用することの限界を指摘したうえで、政策評価論を新しい独自のディシプリンとし、公共政策システム全体をよりよい方向に導くという公共政策学の発想に基づき評価制度を設計し運用すること、そのための評価教育を行うことが必要であることを指摘した。

報告後、コメンテータの西山慶司会員（法政大学特定課題研究所客員研究員）が論点整理を行い、質疑に入った。質疑では、まず評価教育のタイミングについてフロアを交えて議論し、その後他の論点について議論を深めた。参加した会員の多くは評価教育に携わっており、実務の悩みに言及した率直な意見も出され、今後の評価教育の研究と向上に有意義なセッションであったと考える。

共通論題セッション4「自治体評価のキャパシティ・ディベロップメント」

座長 高千穂 安長（玉川大学）

討論者 山谷 清志（同志社大学）

自治体に政策評価が導入されて10年が経過した。多くの自治体で政策評価の実施が日常化しているが、県民、市民など住民からの要望や各種アクターは増加しているにも関わらず、行政改革という外の環境変化から、人、もの、金が毎年のように減少する事態となっているため、政策評価に対して閉塞感が拡大してきている。

このような状況を踏まえ、今後自治体はどのような方向に進むべきなのか、何が問題なのか、あるいは、誰をどのように変えるための評価なのかということについて4名が発表を行い、議論の糸口を提供した。

第1報告者の岩渕公二氏（参議院）からは、地域経営の構造と各アクターの関心から検討を加え、第三者を活用したエンパワーメント機能の効果と課題について論じられた。アクターの増加やローカル・マニフェストの普及が内部評価の限界を生み、その解決のために、住民参加、外部評価が重要であることが示された。ついで第2報告者の小野達也氏（鳥取大学）からは、自治体評価の様々な場面で評価指標などの数値データが活用されているが、評価の客観性を高めることなく、むしろ評価の質を損なうような問題も散見されることを問題提起し、事態の改善に向けて補助情報の活用と数値データの取扱の原則に関わるチェックリストの作成・使用が重要であることを示した。また、第3報告者の鎌田徳幸氏（岩手県庁）からは、岩手県の事例を題材にして、評価を自治体内部および地域社会に定着させるために、評価システムと予算編成システムとの関係の解決、評価に対する職員の意識啓発、地域社会への働きかけが重要であることを示した。最後に第4報告者の高千穂安長氏（玉川大学）からは、川崎市の事例を元に、評価結果の説明責任を果たす段階での多くの問題の存在とその類型および解決のためには、問題所在の認知、組織としての対応と市民(世論)の働きかけが重要であることが示された。

これらの報告に対して討論者の山谷清志氏（同志社大学）から、次の3つの質問提供という形でコメントが寄せられた。1. 評価を既存の枠組みをある意味、壊す形で導入するのはキャパシティ・ビルディング(CD)ではないのではないか、2. 一般企業に公表を前提とする財務会計と公表を前提としない管理会計があ

るが、評価もそのような区分けが必要なのではないか、3. 例えば国の補助事業などは補助金の多寡により事業の進捗は異なるため、評価は必要ないのではないか。

これらについて、発表者からコメントが寄せられ、さらに会場からは政策評価導入のいきさつや3つの質問に対する意見が出され、今後とも自治体政策評価のCD推進の必要性や多くの課題の存在が確認されるなど本セッションの成果は大きかった。

共通論題セッション5「介護サービスの評価」

座長 大島 巖（日本社会事業大学）

座長 平岡 公二（お茶の水女子大学）

少子高齢化社会の進行にともなって、介護サービスに対する社会的コストが増大し、財政負担の引き締めが行われる一方で、質の高い介護サービスの保証が重要な課題になっている。介護報酬不正受給問題が社会的に大きくクローズアップされた背景の一つには、介護サービス、とりわけ訪問介護サービスの収益性の低さが指摘されている。

今日、介護サービスの質を維持・向上させるための評価方法を発展させるとともに、一定のサービスの質を確保し、効果的で効率性の高い介護サービス提供のあり方を求める評価研究が必要とされている。この共通論題では、介護サービス評価の第一線に関わる3名の報告者から、介護サービスの質を高める評価の取り組みについて報告頂いた。

まず、第一報告者である筒井孝子氏には、介護サービスの適正化という観点から、認定データによる給付実態分析の結果にもとづいて介護サービス評価の方向性を議論して頂いた。不適切給付事例の根拠をお示し頂くとともに、事例的に検討することの意義についてご指摘頂いた。また、認定データ、給付データを各自治体が持つがそれを十分に活用していない現状が指摘された。

続いて第二報告者の杉澤秀博氏には、アウトカム指標を用いた介護保険導入前後のパネルデータの分析から介護保険の効果について分析して頂いた。討論の中で、バーンアウトが上昇する理由について、また、関連付けサンプルにすることが死亡者が多い中で可能かどうか、家族構成別のバーンアウトの違いについて議論がされた。

さらに第三報告者の菊澤佐江子氏には、軽度の要介護者に対する介護保険の自立支援効果に焦点を当てて、報告して頂いた。中でも、サービス利用のパターンや利用サービスの効果に対する寄与の分析からサービスの質評価への示唆を頂いた。討論では、従属変数に死亡を入れることの問題（死亡の防止が介護サービスのゴールではないこと）、独立変数として施設利用を入れることの問題、4ヶ月ごとに評価する意味（要介護認定は1年置きなのでその単位でよいのでは）などが議論された。

全体討論では、①インパクト評価としての技術的問題、②サービスの質向上を図る上でのアウトカム評価の活用、③評価主体の位置づけについて議論された。

①インパクト評価としての技術的問題については、介護サービスを利用する特性を持つ人が、「維持」を容易にする特性を持つ可能性があり、因果関係はイベントヒストリー法では解明できないことが議論された。②サービスの質向上に関わるアウトカム評価の役割については、認定データ活用活用の可能性と、アメリカのNursing Home Compareのようにアウトカム評価を用いることの可能性が議論された。③評価主体が、民間・関係団体ではなく、自治体であることについては、民間が主導して質を高めるサービス評価のと、日本において公的規制の制約が大きい点、事業所単位、個人単位で競争原理を導入する必要性について議論された。

介護サービス評価は、今回初めて取り上げたテーマであり、今後これら議論を継続すべきことが議論された。（コーディネータ：大島巖）

共通論セッション6『『地域おこし』の評価』

座長 林 薫（文教大学）

討論者 梅田 次郎（日本能率協会コンサルティング）

開発途上国で村落レベルの生計向上のプログラムが実施、試行されている一方で、日本では都市と地方の格差、高齢化などが深刻な問題になり、地方の振興が叫ばれ、さまざまな「地域おこし」の取り組みがなされるようになってきている。両者には地元のリソースの活用による所得の向上、環境の保全、社会関係資本の強化、女性や高齢者などのエンパワメントなど共通点が多いことから、日本の「地域おこし」を国際協力のプログラムやプロジェクトの中に取り入れる動きも多くなってきている。一方で、日本の経験が途上国に伝えられる際に双方の特有のコンテクストを無視して「形」を模倣になっているのではないか、日本自身が地域おこしの経験を正確に理解し知識として十分に体系化していないのではなかという批判もある。このような批判に応えるためにも地域おこしの評価手法を発展させ、評価の事例を積み重ねることが必要である。

本セッションはこのような問題意識から企画されたもので、まずナオミ ステニング・三好皓一（立命館アジア太平洋大学）による第1報告の“Evaluating Community Capacity Development: The Case of Oyama-cho（地域社会の能力開発の評価：大山町の事例）”では大分県（旧）大山町の事例としてとりあげ、コミュニティーの能力開発（向上）がどのように政策決定の構造にどのような変化を与えてきたかについて考察し、能力の開発によって政策決定の構造もより複雑かつ高度になってきていることが示された。田名邊賢治・小林優輔（総務省行政評価局）による第2報告の「地域おこしの評価方法－非市場財を用いた評価方法－」では、地域おこしの評価は確かに行われてはいるが、基本的には、必要性の理由や事業費で行ったことを定性的に評価しており、「必要だから、必要」といった評価やイベント等の来場者数、開催数等を効果として評価しているものが殆どであるとして、より定量的な評価を目指すべく、近年、社会資本や環境等で用いられている非市場財を用いたさまざまな定量的な評価方法の評価方法について検討を行い、その適応可能性についての報告を行った。

続いて討論者（梅田次郎：（株）日本能率協会コンサルティング）により、地域おこしの運動が行政に取り込まれることによって本来の意義が見失われていることについて問題提起が行われた。この議論を受けて会場から、果たして定量的評価が地域おこしになじむかどうか、（旧）大山町のケースなどは高度経済成長の時期にあたっているという条件を無視すべきではないことなど質問や意見が出された。

本セッションは、地域おこしを評価する場合の評価設問はどうあるべきか、キャパシティー・ディベロップメントの評価をどのように行うべきか等の今後検討を継続すべきアジェンダを引き出せたこと、評価学会での柱である国内の行政評価と国際協力の評価の双方の専門家が議論に参加することによって、両分野の橋渡しを行ったことなどの意義があった。

シンポジウム報告

シンポジウム「中期目標・中期計画の評価」

司会者 松本 哲男（名古屋大学）

中期目標・中期計画に基づくJICAの外部評価は2007年3月に行われ、2008年には国立大学法人の評価が予定されている。そこで本シンポジウムでは、国立大学法人、独立行政法人の中期目標・中期計画を巡る問題を取り上げ、中期計画の評価が次期の中期計画や運営費交付金にどう活かされるべきかについて4名のパネリストを招聘し、討論を行った。

西井知紀氏（文部科学省高等教育局）は、各大学の中期目標の達成状況を、教育研究の質の向上、業務運営の改善および効率化、業務内容の改善、自己点検評価・情報提供などの点から5段階評価する。評価は各大学に改善手当てをしていただくためにするとの見解を述べた。伊藤直樹氏（外務省国際協力局）は、去る3月にJICAの評価を効果的・効率的運営の追求、高い透明性の確保を重点に行い、評価結果に基づく指摘を行ったことを報告した。課題として、目標値の妥当性、中長期的な視野で独法の効率化を図る視点を持つことの重要性和評価が効率化を迫る可能性、評価のための評価の可能性、中期目標・中期計画の策定にかかる多大な行政コストをあげた。栗本英和氏（名古屋大学評価企画室）は「税を使う」から「税をもらう」という大学の意識変革の必要性とともに中期目標の先にあるのは国立大学法人の民営化あるいは国策の実施機関化への懸念、また、現在、評価されるための準備をしているが、そもそも、評価方法を想定して目標や計画を作成していない等法人化時点と評価時との差異があること、評価のための評価、評価の作業化、大学のPDCA（Plan-Do-Check-Act）能力への懸念を示した。山谷清志氏（同志社大学、総務省政策評価・独立行政法人評価委員）からは、「評価の諸形態と組織の対応」として評価の定義、課題、評価から派生する関連業務、数値指標の問題点などの説明があり、評価が最も難しいのが外務関係と大学の先生であり、独法評価はアウトプットであるべきがアウトカムになっているとの指摘があった。最後に討論者の牟田博光氏（東京工業大学理事・副学長）から中期目標・中期計画の評価からは悲しみの未来しかないのではないかとのコメントがあった。

その後、討論に移り、評価結果に基づく次期の運営費交付金配分の問題、評価と大学のランク付け、人事評価を行っている大学の高評価と誘導の問題、民営化等が議論された。特に高評価を受けたからといって予算が増加するのではなく、毎年減っていく決まった額の中での取り合いで疲弊した大学が若者にとって魅力を感じなくなることへの危機感が強く意識されたシンポジウムであった。

自由論題セッション報告

自由論題セッション1「市民社会」

座長 入山 映（サイバー大学）

多義的に使用される「市民社会」というコトバであるが、このセッションでは企業でもなく公的セクターでもない領域、というほどの漠然とした意味で捉える。そして、その市民社会が一方では公的セクターと、他方では企業との間に相互にオーバーラップする領域に着目して、その場面で評価の持つ意味、さらにはそこで重視されるべき要素、といった点に光を当ててみようとした。コメンテーターを置かなかったため三人の発表者が一人30分の発表と10分の質疑応答という比較的余裕を持った時間設定が可能になった。公的セクターとの関わりについては、“The Japanese Red Cross Society's Organization; its Independence or Conformity with Government Structure”（湯井）は日本赤十字社発足時点における皇室との特殊な関係、さらには都道府県行政機構との癒着関係が今日まで持ち越され、市民社会組織としての独立性、あるいはガバナンスが必ずしも確立されていないという問題点を指摘した。フロアからは、特別な法律を持って数を限って設立されるという「特殊法人」性への言及がなく、行政区画との二重構造についてもその解決策、あるいは解消の必要性の有無についての言及がない点についての指摘があった。また「市民社会による日本の対アフリカ政策評価のあり方に関する検討（Ⅱ）」（石田）は昨年の全国大会に続き、TICAD市民社会フォーラム（TCSF）によるアフリカ四カ国の現地NGOによる日本ODAの評価を中心にODAと市民社会の関わり方を分析した。日本ODAにおけるNGO活用については、質量ともに望むべき点が多いが、就中単年度主義と証憑主義については再考を要するのではないかと指摘した。フロアもNGO参加の現実性と阻害要因については関心を示したが、来るべきTICADⅣに向けての市民社会側からの具体的提言内容について、日本のNGOとアフリカのNGOそれぞれの声がいかにして反映されうるのかは残念ながら問題提起にとどまった。他方、企業との関わりについては「新傾向のCSR報告書の評価に関する報告」（井上）が、この境界領域の代表的産物であるCSRについて、その報告書がこれまでの網羅的記載から、めりはりのきいた内容に変化しつつあることを指摘し、それが企業のCSRへの取り組みの姿勢の変化の可能性があることを示唆した。それは同時に報告書が一方的な企業広報ツールではなく、ステークホルダーとのコミュニケーション・ツールとなりうるもの、と考えられた。

自由論題セッション2「政策評価」

座長 小野 達也（鳥取大学）

本セッションでは、「政策評価」というタイトルのもと、まさに政策評価という語の多義性を反映して、評価の主体・対象・手法が異なる4本の報告が行われた。

まず、福岡隆宏氏（高崎経済大学大学院）の「自治体政策評価の活用インセンティブ体系と個人レベルの活用に関する考察」が報告された。本報告は、数の上では広く普及した自治体の評価について、自治体の現場では必ずしも有効に機能していないのではないか、という問題意識のもと、深谷市の評価活動における実体験に基づいて、グループ評価の有効性を論じたものである。会場からは、分析枠組みや他自治体の事例などについて、本研究の今後の進展に資するであろう指摘がなされた。

次に、西本哲也氏（同志社大学大学院）の「政策評価における中央地方関係－政府間評価の試み－」が報告された。本報告は、著者が「政府間評価」と呼ぶところの、政策評価における中央と地方の連携について、欧米の実例に目を配りつつ、我が国の業績測定、プログラム評価におけるケースに注目して、連携の必要性を述べたものである。質疑応答においては、補助金との関係、計画の政府間連携との関係、すでにある種の連携が確立している公共事業評価との関係など、分析枠組みに関する質問やコメントが出された。

3番目の報告は、浅野基樹氏（土木研究所）の「冬期道路管理の業績評価とマネジメントについて」であった。本報告は、著者の所属する研究所において、冬期路面管理という具体的な施策の業績測定について、ロジック・モデルを描き、中間アウトカムと最終アウトカムを試行的に測定した報告である。特に外部要因に左右されない中間アウトカムとして著者らが重視する、路面の連続すべり抵抗値は、新たに導入した装置によって測定されている。質疑では、冬期道路における安全性・走行性に関して、勾配を変数として扱う必要性の指摘などがあった。

最後に、田中弥生氏（大学評価・学位授与機構）・工藤泰志氏（言論NPO）の「市場化テストのPDCAサイクルに見る評価手法の考察～マニフェスト評価体系からのアプローチ～」が報告された。本報告は、政策評価制度が評価対象とするところの政策を超えた、府省の上位にある政治的決定に係る政策の評価について、府省レベルの政策評価結果が無視されたともいえるハローワーク市場化テストの顛末を紹介しつつ、国民によるマニフェスト評価によって、このような政策についてPDCAサイクルを回すことの可能性を論じたものである。質疑応答においては、政策や評価主体、評価目的という基本概念の規定に関して、若干噛み合わないきらいがあったが、この問題の重要性もあらためて認識されたところである。

以上4本の報告の構成から、本セッションに参加された方々の立場や関心領域は様々であったと思われるが、各報告の内容とその後の質疑応答は、それぞれの立場や関心領域を超えて示唆に富むものであったと考える。

自由論題セッション3「教育評価」

座長 石田 洋子（国際開発センター）

本セッションは以下の4件の発表が行われ、それぞれのセッションの後に質疑応答が行われた。

「中学校における授業評価研究」（市野緑、牟田博光）では、現在、日本の小中学校において、児童・生徒による授業評価が組織的に組み込まれつつあるが、実施内容や評価の活用方法に関して多くの課題が指摘されていることから、事例研究として公立中学校の教員に対してアンケートを実施し、その結果を分析したところ、授業評価の活用意識や教員間コミュニケーションが低いことが、授業評価の結果が十分に活用されていないことの原因のひとつと考えられることについて発表が行われた。

「高等教育におけるe-Learningの効果に関する研究結果のメタ分析」（金性希、齋藤貴浩）では、過去12年間（1995年から2006年）の高等教育におけるe-Learningに関する研究結果を収集し、e-Learningの効果について総合的分析とメタ分析を行った結果、pre-postの実験設定自体が実験・統制群の実験設定よりも効果サイズを高める影響を持つてしまうこと、そして個別学習に比べ協業学習の方が、効果が大きいと考えられることについて発表が行われた。

「TIMSS2003における理科の学力に関する要因分析」（藤谷元子、牟田博光）では、理科において、「生徒は実験に参加することによって、生徒中心の活動を体得する。従って、生徒中心の授業が行われるようになり、生徒中心の授業は理科への興味・関心を高めることに繋がる」という仮説を立て、ヨルダンの事例に基づいて共分散構造分析を行った結果、仮説は支持され、教師と生徒の信頼関係が、生徒が実験に参加できる環境及び生徒中心の授業が行われる環境を作る一方、生徒の実験参加により、生徒中心の授業が行われ、生徒中心の授業は、理科への興味・関心を高めることに繋がるということが確認されると共に、コンピュ

ータの活用は、生徒の理科の成績に影響を与えられられることについて発表が行われた。

「大学生生活満足度の要因重要度に関する分析」（星野敦子）では、大学生生活満足度改善のための方法として要因重要度を導入した分析を実施し、「カリキュラム」「成績評価」「資格取得」「個別学生支援」「施設設備」「教員」「就職支援」の7つの要因についてコンジョイント分析を行った結果、最も優先度の高い改善要因は「教員」であること、「成績評価」と「カリキュラム」については重要度は高いが学生の選好にばらつきがみられることから、きめ細かな対応が求められること、「就職支援」「学生支援」についてはより充実させ必要があるが、他の要因と比較してそれほど重要性が高いとはいえず、「施設設備」は継続要因であると考えられることについて発表が行われた。

これらの発表に対し、いずれも、興味深い視点からの教育分野における統計を活用した研究発表であるが、事例研究の際のサンプル数が未だ限られていること、サンプルの抽出に際して、首都圏・地方、公立・私立、ジェンダーバランス等の視点や、対象国の教育開発レベル、社会・経済の背景、文化・習慣等が未だ十分に配慮されていないことなどが指摘された。e-learningや理科の学力に関する分析については、実際の教育現場や教育協力への現場での適用の可能性についても質疑があった。また、それぞれの分析の「視点」の選定の仕方によって統計処理の結果が左右されるケースも考えられ、分析結果を、より客観的で効果的な評価につなげるための考察が必要であることが改めて認識された。

自由論題セッション4「行政評価」

座長 宮崎 修二（対日貿易投資交流促進協会）

当セッションでは、行政における評価に関する実践と評価のパフォーマンス分析、さらには米国における事例についての考察・報告が発表された。

まず、「経営の進捗発信による学校の信頼度向上に関する実践報告（石田謙豪）」では、中小規模の公立小学校における学校評価活動を素材に、評価プロセスを学校運営の向上と改善にどのようにつなげていくかに関しての実践報告があった。報告では、学校運営に対する保護者・地域の信頼度を向上させることが重要であり、そのためには学校が当初に立てた目標に対する進捗状況を定期的に発信することや、管理職と各担当教諭の発信内容とを連鎖させていく努力、発信されるデータの客観性・具体性を確保すること等が重要との紹介があった。こうした努力の結果、発表者の勤務する小学校では保護者からのクレームの減少、学校と保護者との意思疎通の向上等が見られたとしている。本報告に対しては、具体的な実践内容に関する質疑が交わされた。

次の「行政評価制度のパフォーマンスの概念とその分析手法に関する考察（田中啓）」では、自治体における行政評価制度の効果と制度導入の意義を明らかにするための方策として、行政評価制度のパフォーマンス分析に着目し、その概念を明らかにしようという試みが示された。考察では、評価のパフォーマンス分析に関するいくつかの先行研究が紹介された後、行政評価制度の具体的内容（軌跡）と評価実施の成果との関連性を検証する考え方や、個別機関の行政評価制度の発展段階を踏まえた分析結果の解釈の重要性等が示された。これに対し、本考察が概念の枠組み設定の初期段階にあることに鑑み、具体的な評価の実態を踏まえた更なる検証が必要である等の指摘があった。

最後に「米国連邦政府におけるProgram Assessment and Rating Tool: PART運用状況について～業績と予算とのリンケージに関する考察～（左近靖博）」では、米国予算管理庁において2004年度の予算編成から導入されたPART(施策の評価と格付けツール)の運用状況が紹介された。本年、米国政府の全プログラムの評価がPARTによって完了したが、良い評価結果のプログラムには予算が増額配分される傾向が確認される一方、その増減額は一律ではなく、プログラムの性質や種別によって予算額の調整と判断が行われているとの結果が紹介された。これに対しては、政府の予算編成と議会の関係等も含め、PARTの具体的な実

施状況についての質疑がなされるとともに、同制度の今後の運用に関する関心が表明された。

自由論題セッション 5 「国際協力1」

座長 牟田 博光（東京工業大学）

本セッションは以下の4件の発表の後、質疑応答があった。

「Metaevaluation of Aid Evaluations」（佐々木亮）は、事前録画したDVDと米国からのスカイプ利用による質疑応答の組み合わせによる発表であった。主要援助機関（国連機関17、国際融資機関7、二国間援助機関26、EU）によって2004年度に実施され、かつホームページで公開されている事後評価報告書1,034件から層化無作為抽出法により102件のサンプルを選定し、このサンプルについてスクリヴァンのKey Evaluation Checklistを改良したチェックリストに基づきメタ評価を実施した結果、ドナーと分野・地域に関連が強いこと等の分析結果が報告された。

『「目標から自由な評価」考察』（藤田伸子）では、スクリヴァンの「目標から自由な評価」の考え方は、目標や計画にとらわれない、自由な評価を可能にするものの、スクリヴァンが想定したような方法で取り入れることが可能なケースは限定されると思われるが、一般の評価にどのように取り入れることができるかを考察してみると、ある種の発想の転換が促されることが報告された。

「参加型評価による援助パラドックスの解決」（田中博）では、利害関係者の相互理解に重点をおいた参加型モニタリング評価である学び合い評価が提案され、それが援助パラドックス解決に寄与する事が主張された。援助パラドックスとは、利害関係者の不均等な関係から生じる逆説的状況であり、プロジェクト成功を阻害するものであるが、ネパールとフィリピンにおける調査によれば、参加型評価がパラドックスの問題点を軽減する事が報告された。また、手法面として、KJ法や参加型ビデオの活用が有効であることが報告された。

「大学による国際協力機関との連携について―事業受託の問題点とその解決に向けて―」（杉本充邦、松本哲男）は、日本の大学が国際協力事業を受託していく上で共通して抱えている問題点、措置すべき課題について提示し、先進的な取り組みを行っている大学の連携事例を紹介した。また、それに基づき、組織の整備、諸規定の整備、これら事業に従事する人材の確保、さらには外部資金、競争的資金の獲得促進のため、近年導入されたインセンティブの報奨金制度の事例に基づき、問題点解決の提言が行われた。

これらの発表は研究の中間報告、モデルの紹介、実践事例研究、問題提起など、研究として完成したものではなかったが、それだけに内容が新鮮で、活発な議論がなされ、参加者に多くの情報と刺激をもたらしたことを感じさせられた。

自由論題セッション 6 「国際協力2」

座長 林 薫（文教大学）

本セッションは以下の3件の発表の後、質疑応答があった。

「ベトナムの中小企業支援プロジェクトに関する評価」（菊地正、大平哲）は円借款によりベトナム中央銀行を実施機関として四つの取扱銀行を経由した中小企業向けのツーステップ・ローンの評価結果が報告された。評価はDAC5項目に準拠して行われたが、特にインパクトに関して、円借款の資金比率35%までと50～60%の二つの代表的なグループの比較が行われ、前者の方が成功確率の高いことから、中小企業の着実な経営改善とその促進のためには、自己資金を十分に用意できる前者グループへの融資を増やすべきであるとの見解が示された。さらにフィードバック事項として中小企業の定義の見直し、ハノイ、ハイフ

オン、ホーチミン、ダナン以外の地域の資金需要に応じるために地域金融機関のネットワークを活用すべきことなどのフィードバック事項が報告された。

「ネパール政府による公共事業の評価システム強化と日本の支援のあり方について」（石田洋子、三浦順子）はJICA「ネパール国モニタリング評価システム強化計画」（技術協力）の実施を通じて開発途上国の評価キャパシティー・ディベロップメントの実施を通じて問題点や課題を考察したものである。本事業は開始前に十分なベースライン調査行われ国家計画委員会と各セクター別官庁の間の情報共有が不十分であるなどの問題が指摘されていたため、計画の辞しにあたって体制の構築に重点が置かれた。また、コア・トレーナーの研修を核に、政府部門全体のキャパシティー・ディベロップメントを目指すアプローチが採られた。結果としてコア・トレーナーの指導力が予想以上に効果を上げている一方で、複数ドナーの介入によりモニタリング・評価の異なる定義や手法が混在していること、キャパシティー・ディベロップメントの成果を測るのための指標の改善が必要であることなどが報告された。

“Localization of Evaluation—Experience from Nepal and the Philippines”, (Wachowski Cindy Lyn, Miyoshi, Koichi)は評価の現地化が不可欠な要因であることを力説したもので、事例としては上記のネパールの評価システム強化計画とともにEUの支援によるフィリピン のボホールにおけるMDGに向けた現地評価能力強化プロジェクトを取り上げた。これらの経験から、評価の現地化はリソースやキャパシティーが限れている場合でも可能であること、現地かへの政府上層の政治的なコミットメントが必要であること、ジェンダーに配慮したモニタリングが今後の課題であることなどが報告された。

これらの発表に対し、まずベトナムの関する報告に関しては、ベトナムの全体の経済改革（ドイモイ）の評価とその中での中小企業発展の位置づけ、利益率のデータ等の解釈等をめぐって質疑応答が行われた。2報告でとりあげられたネパールの評価のキャパシティー・ディベロップメントにおいては情報共有が行われのレベル、コア・トレーナーの選抜など、評価情報を共有するメカニズムについて議論が行われた。評価の現地化に関しては、評価キャパシティー・ディベロップメントのためには内発的な動機が不可欠であり、そのためにも現地化が不可欠であることについて異論はなかったが、フィリピンのボホールのケースのようにパイロットプロジェクトとして実施する方法と、ネパールの様に現地指導者を育成する方法の対比について議論が行われた。

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る。最近の活動としては以下の活動がある。

1. 研究受託事業

前年度に引き続き、2007年度も国際協力機構より「事業評価の二次評価」業務を受託した。国際協力機構はすべての技術プロジェクト案件について終了時評価を行っているが、年間約50件とその数が多いところから、内部評価として実施している。しかし、評価内容の透明性、客観性を担保するために、作成された評価報告書を外部有識者事業評価委員会で確認する事になっている。外部有識者事業評価委員会の下に作業部会を設け、計量的な分析を加味して、公正なレビューを行っているが、この作業を学会が請け負った。この結果は2007年度末に公表された国際協力機構「事業評価年次報告書2007」の中に集録されている。

国際協力機構からは2007年度新たに「課題別研修第三者検証委員会作業部会」業務も受託した。国際協力機構が平成20年度以降に更新または新設する課題別研修案件の検討において、第三者検討委員会を設置し、案件内容の妥当性および効率性を検証するにあたり、検証プロセスの効率化と客観性のさらなる担保を図るべく、作業部会で審査原案を作成した。

2007年度はさらに、前年度に引き続き、外務省から、「無償資金協力におけるプロジェクト・レベル事後評価に係る評価調査（外務省による評価の二次評価）」業務を受託した。外務省では2005年度から無償資金協力プロジェクトの事後評価を開始したが、各大使館職員が評価に当たるところから、評価内容の透明性、客観性を担保するために、第三者に評価結果の二次評価を委託したものである。この結果は、2007年度末に公表された。なおこの業務は一般競争入札に参加して受託した。

このような受託事業は学会としての社会的貢献を具体的な形で示すと共に、会員に評価作業を経験してもらう場を提供することにもなる。さらに学会に対しては財政的な貢献もできるなど利点が大きいの。今後とも、このような受託事業を積極的に拡大していきたいと考えている。

なお、学会として事業を受託することが適切であるかどうかの判断基準、及び受託事業の品質確保については学会が定めた「受託評価関連事業実施要領」に基づいている。

2. 顕彰事業

学会賞には論文賞、奨励賞、功績賞の3種がある。

2007年度は学会賞受賞者選考細則第6条により、2007年8月1日付けで学会ホームページ上に学会賞候補者の公募を行い、10月10日まで自薦、他薦を受け付けた。10月12日に学会賞審査委員会を開催した。論文賞、奨励賞、功績賞のそれぞれについて、申し合わせに基づき、功績、論文、発表原稿等公表された資料、アンケート結果、及び、日頃の学会活動などを総合的に検討し、さらに、受賞候補者の意思確認のためメールにて議論を継続し、10月29日、学会賞受賞候補者選考について審査員全員一致で下記の結論に達した。

学会賞受賞者選考細則第7条では、論文賞は、選考当該年度の前年度に発行された本学会誌に掲載された原著論文等から推薦されるものとする、と定められているが、今年度は該当なしとした。

奨励賞は、選考当該年度の前年度に開催された全国大会での発表等から推薦されるものとする、と定められている。第7回全国大会、及び春季第4回全国大会の場で発表された原稿等公表された資料、及び、それぞれの大会でのアンケート結果などを総合的に検討し、受賞候補者として星野敦子、佐々木亮会員を選考した。

功績賞は、前年度までの学会活動への貢献に基づき、推薦されるものとする、と定められているが、今年度は該当なしとした。

これらの選考結果は、企画委員長を通じて理事会に諮られ、原案通り承認された。また、受賞者には、11月10日開催の総会場で、学会から賞状及び副賞が授与された。

企画委員会委員長 牟田 博光

研修委員会

活動方針

研修委員会では、「評価の普及および評価に関する人材育成に寄与すること」を目的として、各種セミナーの開催、研修プログラムの開発、教材の作成、講師の派遣、その他人材育成に関わる活動の実施支援を行うことを基本方針としている。2007年度は以下のような活動を展開した。

活動状況

1. 「評価インターン出前サービス」の企画・実施開始

「評価インターン出前サービス」は、学会の学生会員（大学院生）に対して「実際に評価を経験する」機会を提供すべく、「評価を外部に委託したい」と考えている地方自治体やNGO・NPO、学校等で実際に評価を体験させていただく事業である。今年度は、9月より各インターンによる実質的な調査がスタートした。なお、インターンによる調査結果は11月の全国大会（名古屋大学）の場で報告され、2007年3月には、最終報告書がとりまとめられた。

インターン受入れ先	インターン
(1) JICAシリア国リプロダクティブヘルス強化プロジェクト	松本知子（早稲田大学）、村田維沙（東京工業大学）
(2) 社団法人倉敷青年会議所	高田岳志（立教大学）橋本裕人（早稲田大学）

2. 学生向け評価研修の実施

学生向けに評価理論の基礎、評価倫理、事業・組織評価ならびに学校評価の基礎を内容とした評価研修を以下のとおり実施した。本研修は上記評価インターン制度の事前研修としての位置づけもあり、インターン希望者の参加を義務付けたものである。

- (1) 日 時：2007年6月8日（日）9：30～17：00
- (2) 場 所：東京工業大学
- (3) 参加人数：20名
- (4) 講 師：牟田博光（東京工業大学）、高千穂安長（玉川大学）、源由理子（明治大学）、西出順郎（琉球大学）

研修委員会委員長 西出 順郎

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、学会からのお知らせはホームページに随時掲載します」。

より読みやすく、タイムリーに情報をお伝えすべくホームページリニューアルについては引き続き検討していきます。

3. 今後の方針

広報委員会は、現在5名体制です。委員会を開催し、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、効率性を高めて行く所存です。

広報委員会委員長 高千穂 安長

日本評価学会春季第5回全国大会のご案内

春季第5回全国大会

実行委員会委員長・プログラム委員長 牟田 博光

来たる2008年6月8日（日）、日本評価学会春季第5回全国大会が東京工業大学（東京都目黒区大岡山）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お願い申し上げます。

記

1. 日時： 2008年6月8日（日）
2. 場所： 東京工業大学（東京都目黒区大岡山2-12-1）

以上

特定非営利活動法人日本評価学会定款

第1章 総則

(名称)

第1条 本法人は、特定非営利活動法人日本評価学会と称し、英語名は、The Japan Evaluation Society（略称JES）とする。

(事務所)

第2条 本法人は、主たる事務所を東京都品川区に置く。

(目的)

第3条 本法人は、広く国民一般に対する透明性確保、説明責任遂行のため、公共的な活動やNPO等の社会的な活動に関する評価の重要性が強く認識されていることに鑑み、評価に関する研究及び応用を促進し、評価に関する関連機関等との情報交換を図るとともに、この分野の学問の進歩発展及び評価に携わる人材の育成等に貢献することにより、広く一般市民に対し評価活動の向上と評価の普及に寄与することを目的とする。

(特定非営利活動の種類)

第4条 本法人は、前条の目的を達成するため、次の種類の特定非営利活動を行う。

- (1) 社会教育の推進を図る活動
- (2) 学術、文化、芸術又はスポーツの振興を図る活動
- (3) 国際協力の活動
- (4) 男女共同参画社会の形成の促進を図る活動
- (5) 経済活動の活性化を図る活動

(事業の種類)

第5条 本法人は、第3条の目的を達成するため、特定非営利活動に係る事業として次の事業を行う。

- (1) 評価に関する調査・研究事業
広く一般市民の参加を得て客観性を保持し得る評価手法、評価基準の開発を行う。
- (2) 評価に関する研修事業
評価に関する学術研究会、講演会、国際シンポジウム、研修などを開催し、評価人材の育成に努める。
- (3) 評価に関する交流事業
評価研究者、評価実務者、評価利用者等評価に関連する内外の人々との交流、情報交換の場として国際交流や全国大会を実施する。
- (4) 評価に関する広報・普及事業
評価に関する情報を広く知らしむべく、ホームページの作成、学会誌その他刊行物の発行等、普及啓蒙活動を行う。
- (5) 評価実施事業
公平、中立の観点に立ち、法人として第三者評価を実施する。
- (6) 評価に関する顕彰事業
評価に関する研究および評価活動の普及に貢献した人を広く顕彰する。
- (7) その他、本法人の目的を達成するために必要な事業

第2章 会員

(種別)

- 第6条** 本法人の会員の種別は、次の5種とし、正会員をもって特定非営利活動促進法（以下「法」という。）上の社員とする。
- (1) 正会員 本法人の目的に賛同して入会した個人。
 - (2) 学生会員 本法人の目的に賛同して入会した、原則として大学以上の学生で、学生会員を希望する者。尚学生会員は、卒業時に本人が希望する場合は会長に申し込みをおこなったうえで、正会員となることができる。
 - (3) 賛助会員 本法人の目的に賛同し、その事業を後援する団体。
 - (4) 名誉会員 本法人に功労のあった者及び広く評価分野に関連ある分野における学識経験者で理事会の推薦に基づき総会の承認を経た個人。
 - (5) 准会員 本法人の目的に賛同して入会した個人で、最大3年間を期限に本法人の活動への参加を希望する者。3年以上の継続を本人が希望する場合は会長に申し込みをおこなったうえで、正会員となることができる。

(入会)

- 第7条** 会員の入会について、特に条件は定めない。
- 2 正会員、学生会員、賛助会員、准会員として入会しようとする者は、会長が別に定める入会申込書により会長に申し込むものとする。また、名誉会員として推薦された者は、入会の手続きを要せず、本人の承諾をもって会員となるものとする。
 - 3 理事会は、前項の申し込みがあったとき、正当な理由がない限り、入会を認めなければならない。
 - 4 会長は、第2項のものの入会を認めないときは、速やかに理由を付した書面をもって本人にその旨を通知しなければならない。

(会費)

- 第8条** 正会員、学生会員、賛助会員、准会員は、総会において別に定める会費を納入しなければならない。

(会員の資格の喪失)

- 第9条** 会員が次の各号の一に該当する場合には、その資格を喪失する。
- (1) 退会届の提出をしたとき。
 - (2) 本人が死亡し、若しくは失踪宣告を受け、又は会員である団体が消滅したとき。
 - (3) 継続して2年以上会費を滞納したとき。
 - (4) 除名されたとき。

(退会)

- 第10条** 会員は、会長が別に定める退会届を会長に提出して、任意に退会することができる。

(除名)

- 第11条** 会員が次の各号の一に該当する場合には、総会の議決により、これを除名することができる。
- (1) この定款に違反したとき。
 - (2) 本法人の名誉を毀損し又は設立の趣旨に反する行為を行なったことにより、総会において、

除名すべきものと認められた場合

- 2 前項の規定により会員を除名しようとする場合は、議決の前に当該会員に弁明の機会を与えなければならない。

(抛出金品の不返還)

第12条 本法人は、会員がその資格を喪失しても、既に納入した会費その他の抛出金品は返還しない。

第3章 役員

(種別及び定数)

第13条 本法人には、次の役員を置く。

- (1) 理事 20名以上25名以内
 - (2) 監事 1名以上2名以内
- 2 理事のうち1人を会長とし、若干名を副会長とする。

(選任等)

第14条 理事及び監事は、総会において、正会員の中から選任する。

- 2 会長及び副会長は、理事会において、理事の互選により定める。
- 3 役員のうちには、それぞれの役員について、その配偶者若しくは三親等以内の親族が一人を超えて含まれ、又は当該役員並びにその配偶者及び三親等以内の親族が役員の総数の3分の1を超えて含まれることになってはならない。
- 4 法第20条各号のいずれかに該当する者は、本法人の役員になることができない。
- 5 理事及び監事は、相互に兼ねることができず、本法人の職員を兼ねてはならない。

(職務)

第15条 会長は、本法人を代表し、会務を統括する。

- 2 副会長は、会長を補佐し、会長に事故があるとき又は欠けたときは、会長があらかじめ指名した順序で、その職務を代行する。
- 3 理事は理事会を構成し、この定款の定め及び総会または理事会の議決に基づき、会務の執行を決定する。
- 4 監事は、次に掲げる職務を行う。
 - (1) 理事の会務執行の状況を監査すること。
 - (2) 本法人の財産の状況を監査すること。
 - (3) 前2号の規定による監査の結果、本法人の業務又は財産に関し不正の行為又は法令若しくは定款に違反する重大な事実があることを発見した場合には、これを総会又は所轄庁に報告すること。
 - (4) 前号の報告をするために必要がある場合には、総会を招集すること。
 - (5) 理事の会務執行の状況又はこの法人の財産の状況について、理事に意見を述べること。

(任期等)

第16条 役員の任期は、2年とする。ただし再任を妨げない。

- 2 補欠又は増員により就任した役員の任期は、前項本文の規定にかかわらず、前任者又は現任者

の残存期間とする。

- 3 役員は、辞任または任期満了後においても、後任者が就任するまでは、その職務を行わなければならない。
- 4 前項の規定にかかわらず、後任の役員が選任されていない場合には、任期の末日後最初の総会が終結するまでその任期を伸長する。

(欠員補充)

第17条 理事又は監事のうち、その定数の3分の1を超える者が欠けたときは遅滞なくこれを補充しなければならない。

(解任)

第18条 役員が次の各号の一に該当する場合には、理事は理事会の議決により、監事は総会の議決により、これを解任することができる。

- (1) 心身の故障のため、職務の遂行の堪えないと認められるとき。
- (2) 職務上の義務違反その他役員としてふさわしくない行為があったとき。
- 2 前項の規定により役員を解任しようとする場合は、議決の前に当該役員に弁明の機会を与えなければならない。

(費用弁償)

第19条 役員には、その職務を執行するために要した費用を弁償することができる。

- 2 前1項に関し必要な事項は、総会の議決を経て、会長が別に定める。

(顧問)

第20条 本法人に顧問若干名を置くことができる。

- 2 顧問は、会長が理事会の同意を得て委嘱する。
- 3 顧問は、本法人の運営に関し、会長の諮問に答え、又は意見を述べることができる。
- 4 顧問の任期については、第16条第1項の規定を準用する。

第4章 会議

(種別)

第21条 本法人の会議は、総会および理事会の2種とする。

- 2 総会は、通常総会及び臨時総会とする。

(総会の構成)

第22条 総会は、正会員をもって構成する。

(総会の権能)

第23条 総会は、以下の事項について議決する。

- (1) 定款の変更
- (2) 解散及び合併
- (3) 事業計画及び収支予算
- (4) 会員の除名
- (5) 事業報告及び収支決算

- (6) 役員の選任又は監事の解任、役員の職務
- (7) 会費
- (8) 解散における残余財産の帰属先
- (9) その他運営に関する重要事項

(総会の開催)

第24条 通常総会は、毎年1回開催する。

2 臨時総会は、次に掲げる場合に開催する。

- (1) 理事会が必要と認め、招集の請求をしたとき。
- (2) 正会員総数の5分の1以上から会議の目的を記載した書面により招集の請求があったとき。
- (3) 監事が第15条第4項第4号の規定に基づいて招集するとき。

(総会の招集)

第25条 総会は、前条第2項第3号の場合を除いて、会長が招集する。

2 会長は、前条第2項第1号及び第2号の規定による請求があったときは、その日から30日以内に臨時総会を招集しなければならない。

3 総会を招集する場合には会議の日時、場所、目的及び審議事項を記載した書面により、開催の日の少なくとも5日前までに通知しなければならない。

(総会の議長)

第26条 総会の議長は、会長をもってこれにあてる。

(総会の定足数)

第27条 総会は、正会員総数の2分の1以上の出席が無ければ開催することができない。

(総会の議決)

第28条 総会における議決事項は、第25条第3項の規定によってあらかじめ通知した事項とする。

2 総会の議事は、この定款に規定するもののほか、出席した正会員の過半数をもって決し、可否同数のときは、議長の決するところによる。

(総会での表決権等)

第29条 各正会員の表決権は平等なものとする。

2 やむを得ない理由により総会に出席できない正会員は、あらかじめ通知された事項について、書面をもって表決し、又は他の正会員を代理人として表決を委任することができる。

3 前項の代理人は、代表権を証する書面を総会毎に議長に提出しなければならない。

4 第1項の場合において、正会員は、表決内容等について、総会の議長に一任することができる。

5 第1項及び前項の規定により、表決権を行使する場合は、当該正会員は総会に出席したものとみなす。

6 総会の議決について、特別の利害関係を有する正会員は、その議事の議決に加わることはできない。

(総会の議事録)

第30条 総会の議事については、次の事項を記載した議事録を作成しなければならない。

- (1) 日時及び場所
 - (2) 正会員総数及び出席者数（書面表決者又は表決委任者がある場合にあっては、その数を付記すること。）
 - (3) 審議事項
 - (4) 議事の経過の概要及び議決の結果
 - (5) 議事録署名人の選任に関する事項
- 2 議事録には、議長及び総会において選任された議事録署名人2名が、記名押印又は署名しなければならない。

(理事会の構成)

第31条 理事会は、理事をもって構成する。ただし、監事、顧問、幹事ならびに第53条第2項に定める事務局長は、理事会に出席し意見を述べることができる。

(理事会の権能)

第32条 理事会は、この定款に別に定める事項のほか、次の事項を議決する。

- (1) 総会に付議すべき事項
- (2) 総会の議決した事項の執行に関する事項
- (3) その他総会の議決を要しない会務の執行に関する事項

(理事会の開催)

第33条 定例理事会は、毎年3回開催する。

- 2 前項にかかわらず次に掲げる場合には、臨時理事会を開催することができる。
 - (1) 会長が必要と認めたとき。
 - (2) 理事総数の2分の1以上から理事会の目的である事項を記載した書面により招集の請求があったとき。
- 3 理事会は、理事現在数の3分の1以上の出席がなければ、開会することができない。

(理事会の招集)

第34条 理事会は、会長が招集する。

- 2 会長は、前条第2項第2号の場合にはその日から20日以内に理事会を招集しなければならない。
- 3 会長は、緊急に理事会を招集する必要がある場合において、やむを得ない事情によりこれを開催できないときには、理事の承諾を得て、書面により議決を得ることができる。この場合、理事会は開催されたものとみなす。
- 4 理事会を招集するときは、会議の日時、場所、目的及び審議事項を記載した書面により、開催の日の少なくとも5日前までに通知しなければならない。

(理事会の議長)

第35条 理事会の議長は、会長がこれにあたる。

(理事会の議決)

第36条 理事会における議決事項は第34条第4項の規定によってあらかじめ通知した事項とする。

- 2 理事会の議決は、出席理事の過半数の同意でこれを決し、可否同数のときは、議長の決するところによる。

(理事会の表決権等)

第37条 各理事の表決権は、平等なものとする。

- 2 やむを得ない理由のため、理事会に出席できない理事は、あらかじめ通知された事項について書面又は代理人をもって表決することができる。
- 3 前項の場合において、理事は、表決内容等について理事会の議長に一任することができる。
- 4 第1項及び前項の規定により、表決権を行使する場合は、当該理事は理事会に出席したものとみなす。
- 5 理事会の議決について、特別の利害関係を有する理事は、その議事の議決に加わることができない。

(理事会の議事録)

第38条 理事会の議事については、次の事項を記載した議事録を作成しなければならない。

- (1) 日時及び場所
- (2) 理事総数、出席者数及び出席者氏名（書面表決者にあつては、その旨を付記すること。）
- (3) 審議事項
- (4) 議事の経過の概要及び議決の結果
- (5) 議事録署名人の選任に関する事項
- 2 議事録には、議長及び総会において選任された議事録署名人2名が、記名押印又は署名しなければならない。

第5章 資産**(資産の構成)**

第39条 本法人の資産は、次の各号に掲げるものをもって構成する。

- (1) 設立当初の財産目録に記載された財産
- (2) 会費
- (3) 寄付金品
- (4) 財産から生ずる収入
- (5) 事業に伴う収入
- (6) その他の収入

(資産の管理)

第40条 本法人の資産は、会長が管理し、その方法は理事会の議決による。

- 2 この法人の資産は特定非営利活動に係る事業に関する資産とする。

第6章 会計**(会計の原則)**

第41条 本法人の会計は、法第27条各号に掲げる原則に従って行わなければならない。

(会計区分)

第42条 この法人の会計は、特定非営利活動に係る事業会計とする。

(事業年度)

第43条 本法人の事業年度は、毎年10月1日に始まり翌年9月30日に終わる。

(事業計画及び予算)

第44条 本法人の事業計画書及び収支予算書は、会長が作成し、理事会の議決を経た後、毎事業年度の開始前に総会の議決を得なければならない。ただし、やむを得ない事情により、当該事業年度開始前に総会を開催できない場合にあっては、理事会の議決によることを妨げない。この場合、当該事業年度の開始の日から90日以内に総会の議決を得るものとする。

2 会長は、前項の事業計画書及び収支予算書を変更しようとするときは、理事会の議決を得なければならない。

(暫定予算)

第45条 第43条の規定にかかわらず、やむを得ない理由により予算が成立しないときは、会長は、理事会の議決を経て、予算成立の日まで前事業年度の予算に準じ収入支出することができる。

2 前項の収入支出は、新たに成立した予算の収入支出とみなす。

(予備費)

第46条 予算超過又は予算外の支出に充てるため、予算中に予備費を設けることができる。

2 予備費を使用するときは、理事会の議決を経なければならない。

(事業報告及び決算)

第47条 本法人の事業報告書、財産目録、貸借対照表及び収支決算書等決算に関する書類は、会長が毎事業年度終了後遅滞なくこれを作成し、監事の監査を受け、理事会の議決を経た後、当該事業年度終了後90日以内に総会の決議を得なければならない。

2 決算上剰余金を生じたときは、次事業年度に繰り越すものとする。

(臨機の措置)

第48条 予算をもって定めるもののほか、借入金の借入れその他新たな義務の負担をし、又は権利の放棄をしようとするときは、理事会の議決を経なければならない。

第7章 定款の変更、解散及び合併**(定款の変更)**

第49条 本法人が定款を変更しようとするときは、総会に出席した正会員の4分の3以上の多数による議決を経、かつ、法第25条第3項に規定する軽微な事項を除いて所轄庁の認証を得なければならない。

(解散)

第50条 この法人は、次に掲げる事由により解散する。

- (1) 総会の決議
- (2) 目的とする特定非営利活動に係る事業の成功の不能
- (3) 正会員の欠亡

- (4) 合併
- (5) 破産
- (6) 所轄庁による設立の認証の取消し
- 2 前項第1号の事由により本法人が解散するときは、正会員総数の4分の3以上の承諾を得なければならない。
- 3 第1項第2号の事由により解散するときは、所轄庁の認定を得なければならない。

(残余財産の帰属)

第51条 本法人が解散（合併又は破産による解散を除く。）したときに残存する財産は、法第11条第3項に掲げる者のうち、総会において議決した者に譲渡するものとする。

(合併)

第52条 本法人が合併しようとするときは、総会において正会員総数の4分の3以上の議決を経、かつ所轄庁の認証を得なければならない。

第8章 公告

(公告の方法)

第53条 本法人の公告は、この法人の掲示場に提示するとともに、官報に掲載して行う。

第9章 事務局

(事務局の設置)

第54条 本法人に、本法人の事務を処理するため、事務局を設置する。

- 2 事務局には、理事会の同意を得て会長が委嘱する事務局長を置く。
- 3 事務局組織、運営、職員及びその他必要な事項は理事会の議決を経て別に定める。

第10章 雑則

(細則)

第55条 この定款の施行について必要な細則は、理事会の議決を経て会長がこれを定める。

(委員会及び分科会)

第56条 本法人は、事業の円滑な遂行を図るため、委員会及び分科会を設けることができる。

- 2 委員会及び分科会は、その目的とする事項について、調査及び研究し、または審議する。
- 3 委員会及び分科会の組織、構成及び運営その他必要な事項は、理事会の議決を経て、別に定める。

(支部)

第57条 本法人は、事業の円滑な遂行を図るため、総会の議決に基づいて支部を置くことができる。

- 2 支部には、理事会の同意を得て会長が委嘱する支部長を置く。
- 3 支部の位置、組織、運営その他必要な事項は、理事会の議決を経て、別に定める。

附則

- 1 この定款は、本法人の成立の日から施行する。
- 2 本法人の設立当初の役員は、別表の通りとする。
- 3 本法人の設立当初の役員の任期は、第16条第1項の規定にかかわらず、平成16年度の事業報告書及び収支決算の決議をする総会の日までとする。
- 4 本法人の設立当初の事業年度は、第43条の規定にかかわらず、本法人成立の日から平成16年9月30日までとする。
- 5 本法人の設立初年度の事業計画及び収支予算は、第44条の規定にかかわらず、設立総会の定めるところによる。
- 6 本法人の設立当初の会費は、第8条の規定に関わらず、次に掲げる額とする。

正会員	15,000円
学生会員	5,000円
賛助会員	1口100,000円（1口以上）

【別表】

設立当初の役員

役 職 名	氏 名
会 長	宮内 義彦
副 会 長	廣野 良吉
副 会 長	牟田 博光
理 事	竹内 正興
理 事	荒木 光弥
理 事	梅内 拓生
理 事	喜多 悦子
理 事	木村 孟
理 事	熊岡 路矢
理 事	坂元 浩一
理 事	白須 孝
理 事	杉下 恒夫
理 事	鈴木 絲子
理 事	高千穂安長
理 事	高橋 一生
理 事	長尾 真文
理 事	古川 俊一
理 事	三好 皓一
理 事	森 茂子
理 事	弓削 昭子
監 事	高瀬 国雄
監 事	松下 弘

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

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。
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 日立ソフトタワーB 22F
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@idcj.or.jp

『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1 ○○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 ○○

--

Note:
Source:

Table 1 Number of Accidents in the State of ○○

--

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 6th edition*. Beverly Hills, Calif: Sage Publications.

英語雑誌論文：著者（発行年）. 題名. 雑誌名, 巻（号）, 頁－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.

3名以上の著者による英語文献：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）おとして記述する。

『日本評価研究』 査読要領

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。

～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 日立ソフトタワーB 22F
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8巻第1号

2008年3月26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
日立ソフトタワーB 22F
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. 8, No. 1, March 2008

CONTENTS

Special Issue: A Decade of Evaluation Practice in the Japanese Local Government

Editor's Note Kiyoshi Yamaya

Hardships of the Pioneer :
Introducing the Evaluation Systems in the Mie Prefectural Government Jirou Umeda

Ten Years of Evaluation Practices in the Prefectures of Japan
–The Past and Future of Quantitative Methods– Tatsuya Ono

Performance Measurement in City Governments in Japan:
The Current State Ten Years after Its Introduction Hiraki Tanaka

Public Policy Studies, Evaluation Studies, and the Mode of Evaluation in Japan
Yoshio Kubota

Outside Supporting Services for Local Governments
in Introducing and Stabilizing the Performance Management System Yoshiaki Hoshino

Collaboration in Evaluation Systems of Non-Profit Sector and Local Governments
Koji Iwabuchi

Articles

Examining the Limitation of the Current Policy Evaluation System in Japan:
The Case of Policy Management Problems of the Market Testing System Executed at
HELLO WORK Yayoi Tanaka, Yasushi Kudo

Study on the Municipalities Should Do to Improve Evaluation Ability
What Municipalities Need to do Yasunaga Takachiho

Research Notes

Six Modes of Evaluation Use Go Yoshizawa