ISSN 1346-6151

日本評価研究

Japanese Journal of Evaluation Studies

Vol. 16, No. 1, August 2016

特集：自治体評価の20年

巻頭言　自治体改革と評価　山谷 清志

自治体における業績測定型評価の現状と課題
—20年を経過した都道府県の取り組みの点検結果から—　小野 達也

自治体評価を振り返る
—「活かさず殺さず」の20年　西出 順郎

地方分権改革と財政危機の自治体評価
—20年のレビューから—　山谷 清志

研究ノート

ネットワークによる公共サービスのマネジメント・ガバナンスと業績評価
—欧米の先行研究に見る動向と今後の研究課題—
荒川 潤　玉村 雅敏

春季第12回全国大会開催報告

第16回全国大会開催報告

第17回全国大会のご案内

委員会活動報告

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
山谷　清志（同志社大学）
Kiyoshi YAMAYA

副委員長
Vice-Editor-in-chief
西野　桂子（関西学院大学）
Keiko NISHINO

常任編集委員
Standing Editors
牟田　博光（東京工業大学）
Hiromitsu MUTA

編集委員
Editors
岩渕　公二（NPO政策21）
Koji IWABUCHI

大島　巌（日本社会事業大学）
Iwao OSHIMA

岡本　義朗（新日本有限責任監査法人）
Yoshiaki OKAMOTO

小野　達也（鳥取大学）
Tatsuya ONO

窪田　好男（京都府立大学）
Yoshio KUBOTA

佐々木　亮（国際開発センター）
Ryo SASAKI

渋谷　和久（内閣府）
Kazuhisa SHIBUYA

田中　弥生（大学評価・学位授与機構）
Yayoi TANAKA

西出　順郎（岩手県立大学）
Junro NISHIDE

林　　薫（文教大学）
Kaoru HAYASHI

松岡　俊二（早稲田大学）
Shunji MATSUOKA

源　由理子（明治大学）
Yuriko MINAMOTO

事務局
Office
〒108-0075　東京都港区港南1-6-41 品川クリスタルスクエア 12階
一般財団法人国際開発センター内
特定非営利活動法人日本評価学会
TEL: 03-6718-5931, FAX: 03-6718-1651
E-mail: jes.info@evaluationjp.org

『日本評価研究』は、和文年2回、英文年1回を刊行します。

日 本 評 価 研 究

第16巻　第１号　2016年８月

目　次

特集：自治体評価の20年
山谷 清志
巻頭言　自治体改革と評価…… 1

小野 達也
自治体における業績測定型評価の現状と課題
―20年を経過した都道府県の取り組みの点検結果から―…… 3

西出 順郎
自治体評価を振り返る　―「活かさず殺さず」の20年…… 17

山谷 清志
地方分権改革と財政危機の自治体評価　―20年のレビューから―…… 31

研究ノート
荒川 潤　玉村 雅敏
ネットワークによる公共サービスのマネジメント・ガバナンスと業績評価
―欧米の先行研究に見る動向と今後の研究課題―…… 47

春季第12回全国大会開催報告
開催の報告とお礼…… 63
プログラム詳細…… 64
共通論題セッション報告…… 66
自由論題セッション報告…… 71

第16回全国大会開催報告
開催の報告とお礼…… 74

プログラム詳細……75
共通論題セッション報告……78
自由論題セッション報告……85
シンポジウム報告……89

委員会活動報告
企画委員会……91
国際交流委員会……92
広報委員会……93

第17回全国大会のご案内 ……94

日本評価研究刊行規定……95
日本評価研究投稿規定……97
日本評価研究執筆要領……99
日本評価研究査読要領……102

Publication Policy of the Japanese Journal of Evaluation Studies ……104
Information for Contributors (For English Papers) ……106
Writing Manual of the Japanese Journal of Evaluation Studies (For English Papers) ……108
Referee-Reading Guideline……110

【巻頭言】

「自治体改革と評価」

山谷　清志
同志社大学 政策学部

地方自治体に「評価」が登場してほぼ20年、さまざまな意味でそのあり方を振り返る時期に来ている。日本評価学会は設立当初からこの地方自治体評価を議論対象の一つとしてきたが、国際的に見るとそこには顕著な特徴がある。

第1に、外国の文献では目にしないPDCA（Plan-Do-Check-Action）を多用する傾向である。2003年2月17日、「平成15年第3回経済財政諮問会議」の奥田碩議員（当時トヨタ自動車会長）の発言以来、PDCAは日本国内で急速に広がり、地方自治体も影響を受けた。しかしこのPDCAの考え方は評価理論とは無縁である。なぜなら奥田議員が言うように、PDCAは生産現場の日本型QC活動「カイゼン」にルーツがあり、企業経営や生産現場で重視された日本オリジナルの実践活動で、諸外国の評価実務家や理論家にはなじみがない。そのため、評価の国際会議で言及しても理解が得られない。もちろんcheckはevaluationではない。

第2に、客観性にこだわりすぎるところである。評価の客観性と中立性を求め、第三者評価の制度に熱心なのは、評価プロセスの透明性、公平性を確保するのが目的である。しかし、客観性評価は欧米の評価専門誌に登場しないし、実務でも見られない。逆に、客観性よりはcapacity-building、empowermentのテーマが多く目につく。より多くの人びとを評価プロセスに参加させ、多様な意見を評価に反映させたいのである。「権威ある」有識者の知見を重視する日本の慣行とは、基本的に考え方が違う。日本では評価が客観的、公正、公平であるべきだという信念があり、もし客観性を疑われると困る事態が多いので（卑近な例であれば学校の成績評価や入試）、過剰に意識するのである。他方で、日本の国技である相撲では、時に行司は力士と同じ部屋に所属し、また力士の親方が勝負審判を務める。それについて客観性を問う意見は聞いたことがない。評価においては状況と場面で客観性を使い分けるのが現実的であるが、客観的であるという形式だけが重視されているので、手間がかかる。

特徴の第3は、先のPDCA概念を用いてCとAが重要だと主張する点である。フィードバックが大事だからと言うのはよく分かる。しかし、このフィードバックのために何をするのか、それがはっきりしない。予算編成に反映させたいのはよく分かるが、事業が終わっていたら反映させる予算がない。複数年度の途中で効果が無いと評価されれば次年度の予算の減額も考えられるが、それは事業の継続を保証することになる。廃止ではないからである。もし評価で予算が減額されることが分かっていれば、事業担当者にとって評価はディスインセンティブの仕組みとして働く。地方自治体ではそもそも、政策の仕組みやデザインが悪くてはじめから効果が出ないと分かっている政策、目的が曖昧で達成できるかどうか不明な政策、どう考えても目的を達成できない政策手段が所与になっている政策、全く成果が出ないまま惰性で繰り返されている政策が多い。また、CとAに重大な欠陥がある行政の仕組みもある。こうした状況でフィードバックを求めると、現場の担当者に過大な労苦を強いる。

第4は、用語とそれにまつわる機能に対する理解の特徴である。地方自治体で多用される事務事業評価

日本評価学会『日本評価研究』第16巻第1号、2016年、pp.1-2

はどう見ても「測定（measurement）」で、年度ごとの数字をちりばめた評価表や評価シートは、経済分析やシステム分析のような「分析（analysis）」がない記述（description）が多い。それを評価（evaluation）と呼ぶところにユニークさがある。むしろ小野論文のように「業績測定」であると言ってしまった方が、市民にとって混乱がない。

欧米の評価の歴史をふり返れば、第二次世界大戦前から、社会科学を応用する形で公的機関活動のレビュー、リサーチ、スタディが福祉や教育、就労支援・職業訓練の分野で試みられてきた。これらを評価（evaluation）と呼ぶのは周知のことである。あまりにも普通のことであるため、意識されていない。

続く1950年代のシステム分析、1960年代のエコノミストによる経済分析などが行動科学（Behavioral Science）と結びつき「政策科学（Policy Sciences）」と呼ばれ、「分析の時代」が訪れた。これらの分析が予算編成の実務と結びつき、PPBS（Planning-Programming-Budgeting System）として、政府機関や国際機関、そして日本政府でさかんに用いられた。日本評価学会のシニア会員が経験され、よくご存じである。日本語ではプロジェクト分析、事業評価として知られ、定着もしている。並行して1960年代末から70年代にかけて、アメリカでは福祉プログラムを評価する要請が高まり、プログラム評価（program evaluation）が流行する。いま日本で総合評価と呼ばれる政策評価方式のプロトタイプである。こうした一連の流れを知る経済学関係者はanalysisを理解し、教育学や福祉学の専門家はevaluationを実践している。

1980年代末、日本の行政研究者の間ではイギリスのサッチャーの政府改革に関心が集まり、Next Steps改革が注目される。政府の政策実施機能を外部に切り出すAgency、あるいはExecutive Agencyと呼ばれる仕組みがとくに注目され、それらエージェンシーの活動実績を判断するものさしとしてperformance indicatorが置かれた。このindicatorを使って活動実績を測定するのがperformance measurementである。日本でこれをまねたのが独立行政法人制度であり、また実績評価すなわち業績測定である。この測定、イギリスではoutputの測定であったが、日本では1993年にゴア副大統領が主導したアメリカのReinventing Government改革の影響を受けoutcomeの測定になった。会計学に詳しい人や公認会計士で英米の行政改革を見てきた人はperformance measurementになじみがある。

しかし、これらの研究や実践に無縁の公務員が一番取り組みやすいのは「目標の達成度」を測定する方法である（ただし小野論文が示す手順に従っていない場合が少なくない）。日本で地方自治体の行政評価が業績測定になり、シート記入という公務員がなじんだ方式に落ち着いたのは、効用の有無は別として必然である。もちろん、行政評価がブーム化してシンクタンクやコンサルタントのビジネスチャンスになった時、さまざまな提案がなされた。たとえば自治体経営戦略の見直し、意識改革のためにQC、ISO、カイゼン、市場化の導入であった。

自然大災害、市町村合併、リーマン・ショックの大変動を経験した今、自治体評価を冷静に見直す時期に来ている。自治体評価は縮小簡便化、あるいは積極活用（西出論文）、このどちらに進むのであろうか。どんな方法がベストなのであろうか。鍵を握るのは市民のシチズンシップ、公務員の知的洗練、地方政治家のガバナンス感覚であろう。いずれにも、その背景に「評価文化」があることは間違いない。

【研究論文】

自治体における業績測定型評価の現状と課題
ー20年を経過した都道府県の取り組みの点検結果からー

小野　達也
鳥取大学地域学部
onotatsu@rs.tottori-u.ac.jp

要　約

自治体に最初の行政評価制度が導入されて20年が経過した。その行政評価制度の中心はいわゆる業績測定型の評価である。その手法はシンプルであり、最初に導入した三重県を始め一定の結果が得られたこともあって、数多くの自治体が業績測定型評価に取り組むに至っている。しかし、業績測定型評価の成否は、その基本ツールである評価指標の品質や取り扱い次第である。本研究では、自治体の行政評価において先導的な役割を果たしてきた都道府県の取り組みを点検し、また同様の点検を行った平成19年時点の状況とも比較した結果、改善が見られる面もあるものの依然として多くの課題があることが明らかになった。それらの多くは、政府の目標管理型政策評価など公共部門の業績測定型評価に共通する課題でもある。業績測定型評価が本格的に機能するためには、これらの課題への対策が必要である。

キーワード

業績測定（パフォーマンス・メジャーメント）、都道府県、評価指標、目標達成度

はじめに

平成8年度に三重県が事務事業評価を開始してから、行政評価[1]に取り組む自治体は増加の一途を辿り、総務省の調査（総務省2014）によれば、平成25年には都道府県と市区のうち86%、町村でも35%が行政評価を行っている。近年は評価のルーチン化や形骸化、評価疲れなど様々な問題も論じられているが、平成22年からの3年間をみると、すでに100%近い都道府県・政令市・中核市・特例市の導入率は横ばいまたは微増、その他の市区の導入率は78%から83%へ5ポイント上昇、町村の導入率も30%から35%へ5ポイント上昇している[2]。これは多くの自治体が自らの現状と今後の行財政において行政評価の必要性を認識しているからに他ならないであろう。

自治体の行政評価の多くは業績測定型評価[3]である。三重県の事務事業評価も米国地方政府の業績測定型評価の系譜に属するものであり、評価指標による目標管理が重要な役割を担っていた。後に続く自治体も、評価指標の役割に大小はあっても、大半は業績測定型評価の性格を有するといってよいだろう。

上述の総務省の調査によれば、行政評価導入済みの1,060団体のうち評価指標を用いているのは85%の900団体（都道府県は94%、政令市は100%、市区町村は84%）である。なお、行政評価導入済みの団体のうち総合計画の進行管理に

日本評価学会『日本評価研究』第16巻第1号、2016年、pp.3-16

活用している団体が71%あり、総合計画の進行管理を業績測定型評価によって行うのが典型的な形であると考えられる。

業績測定型評価の成否、すなわち評価が機能するか否かがその基本ツールである評価指標の品質や取り扱い[4]によって大きく左右されることは言うまでもない。一方、自治体に行政評価が導入された初期の段階から一貫して、数値の指標設定、特にアウトカム指標の設定が難しいという声が少なくなく、指標の品質に自信のない自治体が依然として多いことが様々なアンケート結果に示されている（例えば田中2014、佐藤2012）。

本稿では、概念は極めてシンプルでありながら業績測定型評価において決定的な役割を担う評価指標による目標管理について、特にその道具立てとしての品質に焦点を絞り、導入から20年を経過した自治体の現状を点検する。直接点検の対象とするのは、自治体の行政評価において先導的な役割を果たしてきた都道府県である。

筆者は最初の行政評価導入から10年余り経過した平成19年10月にも47都道府県の業績測定型評価の点検を行い、10項目の問題点を指摘した（小野2008）。今回の点検は都道府県の業績測定型評価の品質がこの10年間でどの程度改善されたのかを確認する作業でもある。

前回の点検作業の後、筆者は業績測定型評価における目標設定から達成度評価、さらには評価結果の集約にいたる各段階の方法論上の課題について詳細に検討している（小野2011）。今回はこの結果も踏まえ、都道府県の業績測定型評価の現状をより詳細かつ体系的に点検することとなる。

業績測定の方法自体は市町村や国も共通である。実際、国の業績測定型評価（平成12年度に実施された国の政策評価制度の目標管理型評価及び行政事業レビュー）を対象に筆者が点検した結果、多くの共通した課題が明らかになった（小野2013）。

国レベルの業績測定型評価については、これまで数次にわたり定量評価の改善（定量指標の使用の徹底など）が外部からも求められてきたが、最近も注文が相次ぐとともに政府自らの見直しも進行中である[5]。

さらに近年は、行政評価の制度・システム以外の様々な場面においても業績測定による目標管理が求められるようになっている。国・自治体の様々な政策分野の計画や、自治体の首長が掲げるマニフェスト、平成27年度中に大半の自治体が策定するに至った地方創生の総合戦略におけるKPIなどである。

本稿は今回の点検結果を踏まえ、自治体の業績測定型評価を念頭に課題の解決策を検討するが、それらは業績測定型評価一般にも少なからず該当するものとなろう。

1. 評価指標の取り扱いに係る点検項目

評価指標による目標管理の過程は、評価指標の設定、目標値の設定、目標達成度の把握・評価、目標達成度の評定と集計、という4段階に分けて考えることができる。本稿では全都道府県の業績測定型評価を対象に、この各段階について、以下の計10の項目の点検を行う。なお、これらの項目は平成19年の前回調査と概ね一致する（前回調査では項目③と④は合わせて1つの項目）。

評価指標の設定

①指標の妥当性を示す情報

評価指標に一定の妥当性（測定すべきものを測定していること、すなわちプログラム（以下の議論では政策・施策・事務事業を総称してプログラムと呼ぶ。）の進捗を判断するのに適した事象や状態を測定・把握できること）がなければ、目標管理のすべてが意味を失う。本研究では、指標の妥当性を示す何らかの情報が明らかになっているか否かを点検する。用いられている指標の妥当性自体は直接評価していない[6]。

②フローとストックの区別

評価指標はフロー（一定期間に発生する量）を測る指標、ストック（ある時点に存在する量）を測る指標、それ以外の指標の3種類に分類できる。また「それ以外の指標」も行政の注力の変化と指標値の変化の関係に基づいて「フロー的」

指標と「ストック的」指標に分類することもできる（小野2011）。フロー指標の推移とストック指標の推移を区別することなく直接比較することはできない。

目標値の設定

③設定の根拠

目標の数値を設定するには通常、過年度の実績や他地域の水準、ニーズを示すデータなど、何らかの根拠があるはずである。その説明の有無、仕方を点検する。項目①同様、目標値自体が妥当か否かまでは問わない。

④目標値の性格

目標の意味には様々なものがあるはずである。非常に高く挑戦的な水準、一定の注力を続ければ達成できると見込まれる水準、問題のある現状を改善して達成することが必須の水準など。例えば僅かながら目標未達となった場合、その実績をどのように評価すべきかは当然大きく異なる。目標値の性格に関する情報の有無を点検する。

⑤指標値の増減パターンの想定

過年度の実績をもとに将来の目標値を設定したり、中期的な目標値をもとに各年度の目標値を決めたりする際など、指標値改善の一定ペースを想定する場合がある。このとき「直線」（毎年度一定量の増加など）を想定する場合が多いが、予算額に比例するようなアウトプット指標はともかく、様々な過程を経由して、あるいは複合的な要因の影響を受けて発現するようなアウトカム指標の場合、毎年一定の上昇率であることを想定するのが自然である（小野2011）。

目標達成度の把握・評価

⑥目標達成度の計算式

目標達成度を計算する場合、その計算は実績値を目標値で除す方法（本稿では「水準達成度」と呼ぶ）と基準値から実績値までの変化量を基準値から目標値までの変化量の目標で除す方法（本稿では「変化達成度」と呼ぶ）に二分される[7]。両者の意味は当然異なるが、例えばストック指標の水準達成度の解釈は要注意である。水準達成度の計算は期首の水準を考慮せず、値がゼロの時点から期末までの変化達成度に相当するからである（小野2011）。

⑦例外的な場合の目標達成度

目標達成度を計算する場合、小さい値が望ましく減少を目指す指標の水準達成度、現状維持を目指す場合の変化達成度など、例外的な取り扱いが必要な場合がある（小野2011）。前者のケースで計算式の分子・分母を逆転させ目標値を実績値で除すなど根拠不明な方法が少なからず見られる。

⑧外部要因の把握

業績測定型評価においては、アウトカム指標値の変化に外部要因（評価対象プログラム以外の要因）が及ぼす効果を定量的に把握することはない（基本的に不可能）が、定性的に把握することの意味は大きい。

目標達成度の評定と集計

⑨目標達成度の評定

評価対象のプログラムの進捗状況を5段階の評定などで簡明に示す場合も多い。目標達成度の率の数字を階級に分けてカテゴライズする方法や目標達成度の数字にその他の情報も合わせて複合的に評定する方法、目標達成度の計算はせずに総合的な判断に基づく方法など様々だが、評定の基準は明確でなければならない。

⑩目標達成度の集計・集約

総合計画全体の進捗状況など、評価プログラム全体（ないしその部分集合）の目標達成度を何らかの集計により説明する自治体も多い。⑨の評定結果を評点ごとに数える、評点の平均を計算する、目標達成度の数字を平均するなど様々な方法があるが、何れも評定の結果が比較可能であること（同じ「性格」の数字であること）が前提となる。また評点を平均するためには、評点の数字が本来間隔尺度以上でなければならない。

２．今回の調査の方法

今回の点検は、全都道府県の業績測定型評価を各団体のホームページ上で調査する方法によ

表１　都道府県の業績測定型評価（今回の点検対象とした制度・システム）

都道府県	制度・システム	評価実施年度
北海道	施策評価	26
青森	政策点検	27
岩手	政策評価	26
宮城	施策評価	27
秋田	政策評価	26
山形	総合計画の進行管理（評価・検証）	26
福島	総合計画の進行管理	25
茨城	政策評価 施策・事業評価	25 25
栃木	重点戦略マネジメント	27
群馬	総合計画の目標・指標の達成状況	27
埼玉	政策評価	26
千葉	総合計画の進行管理	26
神奈川	総合計画の点検報告	26
新潟	行政計画推進のための評価 政策分析	27 27
富山	政策評価	26
石川	目標管理型行政経営	26
福井	政策合意の達成状況	27
山梨	行政評価	26
長野	計画進捗管理	26
岐阜	長期構想の長期目標の達成度	−
静岡	施策展開表	27
愛知	管理事業評価	26
三重	施策評価	27
滋賀	総合計画の進行管理（プロジェクトの進捗状況）	26
京都	中期計画の進捗状況（ベンチマークレポート）	26
兵庫	長期ビジョンの点検・評価（全県ビジョンフォローアップ指標）	26
奈良	重点課題に関する評価	26
鳥取	政策項目の進捗状況 工程表	25 26
島根	施策評価	26
岡山	政策評価（暮らしやすさ指標）	26
広島	施策マネジメント（点検・評価）	25
山口	政策評価（プランの進行管理）	24
徳島	主要事業等評価シート	26
香川	施策評価（9 つの指標） 施策評価（100 の目標値）	26 26
愛媛	政策事務事業評価（予算施策評価表） 政策事務事業評価（事務事業評価表） 長期計画の進捗状況評価	26 26 25
高知	計画の進捗管理	26
福岡	事務事業評価 総合計画実施状況（施策目標の進捗状況）	27 26
佐賀	施策評価（実績評価）	27
長崎	政策評価（事業評価）	26
熊本	戦略指標の動向	26
大分	施策評価	26
宮崎	政策評価（「目標値」達成状況の評価）	27
鹿児島	行政評価（施策評価）	26
沖縄	PDCA（実施結果）	26

（注）制度・システムの名称が明確でない場合も多く、筆者がホームページ上の記述から適宜選択した。総合計画の…という名称以外にも総合計画の進行管理手法であるものは多い。
（出所）筆者作成。

り行った（前回の調査と同様）。ホームページを閲覧した期間は平成27年9月2日～9月8日で、その時点で最近の年度の業績測定型評価を点検対象とした。

業績測定型の評価制度・システムを導入している団体について、制度・システムが1つであればそれを取り上げることとし、複数の制度・システムがある場合（例えば総合計画の進行管理において政策－施策－事務事業の複数段階について同一でない方法で評価を行っている場合や総合計画の進行管理と事務事業評価が別の制度となっている場合など）は、業績測定の性格が明確であり、評価情報・データの品質が高いと思われるものを1つ選ぶことを原則とした。ただし、手法に明確な違いがある業績測定型評価が複数あって（例えば施策レベルと事業レベルで評価指標の取り扱いが異なる場合など）、品質に大差がない場合は、それぞれを取り上げることとした。

なお、前回の調査では、各都道府県から最も充実した内容の業績測定型評価を1つだけ選んで点検の対象としており、今回調査の結果との比較には注意が必要である。

今回の調査で点検の対象とした業績測定型評価は44団体の51件である（前回調査では44団体）。それらの一覧を表1に掲げた。なお、東京都・大阪府・和歌山県については、近年行われた業績測定型評価の情報を見つけることができなかった。

上述の総務省調査（調査時点は平成25年10月であり、本研究の調査対象は表1に示す通り、平成25年度実施のものもあるが大半は平成26、27年度実施のものであるため、若干の時点の違いがある）では、評価指標を用いた行政評価を行っている都道府県は44団体であり、今回の点検対象の制度・システムがある44団体と数は一致する。

3. 都道府県の業績測定の現状－今回の調査結果及び前回との比較から

項目ごとの点検結果は以下の通りである。

①指標の妥当性を示す情報

表2に示す通り、プログラムのロジック（資源投入やその手段から成果の発現までの過程の論理的関係）上の何れの段階を評価指標の把握対象としているか、評価指標の基本的な妥当性が十分に示されているような事例はない。評価指標が当該プログラムのアウトカムに関わることが大まかに示されていたり、評価指標の持つ意味に関する説明が多少ともあったりするなど、妥当性に関わる何らかの情報が示されている割合は2/3である。前回調査では44団体のうち「妥当性に関する記述、吟味」が何らかの形であったのが25団体（57%）であったから、状況は若干改善しているともいえるが、大半の評価書や計画書において指標の妥当性に関する記述が不十分または全くないという状態である。

表2　指標の妥当性に関する情報はあるか

分類	件数	割合
ロジック上の位置などが明確である	0	0.0%
プログラムとの関係がある程度わかる	8	15.7%
何らかの情報がある	26	51.0%
説明なし	17	33.3%
計	51	100.0%

（出所）筆者の今回の調査による。

②フローとストックの区別

表3に示す通り、明確に分類しているのはわずか4件（北海道、宮城、愛媛2件）に止まっている（前回調査では3団体）。フローとストックに近い分類として単年度・年間と累計に分類している事例も4件あり、ストック指標と一定期間のフロー指標の累積を合わせて「累計」とすれば（秋田、熊本）、フローとストックの別と同様の分類となる。

表3　フローとストックの区別はあるか

分類	件数	割合
フローとストックを明確に分類	4	7.8%
単年度・年間と累計に分類	4	7.8%
区別せずに混在	43	84.3%
計	51	100.0%

（出所）筆者の今回の調査による。

ただし、フローとストック、単年度と累計の区別は、その取り扱いにおいて両者を区別することで初めて有意義な分類となる。後述するが、2つのタイプの指標を全く区別せずに扱えば、点検項目⑥⑨⑩の段階で少なからぬ問題が発生する。

③目標値設定の根拠

目標値を設定しているのは43団体の50件である。表4に示す通り、目標値の設定根拠が明らかなのは18件であり、半数以上の事例では全く説明がない状態である。

表4　目標値の設定根拠は示されているか

分類	件数	割合
明確に説明	18	36.0%
一定の説明がある	4	8.0%
全く説明なし	28	56.0%
計	50	100.0%

（出所）筆者の今回の調査による。

④目標値の性格

表5に示す通り、ほとんどの場合、目標値の性格に関する情報はない。わずかに岐阜県が目標値の「設定パターン」を6通りに類型化しているほか（ただし、その意義は目標達成状況の解釈や意思決定への反映段階でその分類が活かされているか否か次第である）、福島県が目標値について「原則として、県の施策の努力目標となる数値を記載」としているのみである。

項目③と合わせ、目標値の設定根拠と性格について何らかの説明があるのは23件に留まっており、前回調査において25団体であったことと比べて、状況の改善は見られない。

表5　目標値の性格は明らかか

分類	件数	割合
明確に分類している	1	2.0%
大まかに統一している	1	2.0%
全く情報なし	48	96.0%
計	50	100.0%

（出所）筆者の今回の調査による。

⑤指標値の増減パターンの想定

将来の目標値の設定にあたってベースラインとしてどのような想定を行うかを明らかにしたり、年度ごとの目標値を設定しない場合に、中期的な目標年度の目標値に基づく中間年の評価基準の置き方を明らかにしたりすることは望ましいことである。しかし、そのような場合に直線状の等差的な推移のみを一律に想定してしまっている事例が8件あった。

一方、途中時点での達成見通しを個別の指標ごとに議論したり、目標値の設定根拠として伸び率一定をベースとする指標があったりするなど、直線状の増減パターンに限らない取扱いが確認できた事例も4件あった（岩手、宮城、富山、石川）。前回調査では1例も確認できず、丁寧な取扱いが一部で登場したといえる。

⑥目標達成度の計算式

実績値も目標値も数値であれば、両者から目標達成度を計算することは自然であり、目標達成度を客観的に把握するために必要であるともいえる。今回の調査では目標達成度の計算をしている制度・システムは27件で目標値を設定している事例のうち54%であった。前回の調査では、目標値を設定している44団体のうち24団体（50%）であり、割合は同程度である。

各事例がどのような計算を行っているかを表6に示す。水準達成度または変化達成度の計算をすべての指標について行う事例が多いことがわかる（前回調査では一律の水準達成度が11団体、一律の変化達成度は8団体）。ここで点検すべき

表6　目標達成度の計算方法

分類	件数	割合
目標達成度を計算	27	54.0%
一律に水準達成度を計算	12	24.0%
一律に変化達成度を計算	12	24.0%
水準達成度と変化達成度の使い分け	3	6.0%
水準達成度と変化達成度の併用	0	0.0%
改善幅を計算	1	2.0%
計算せず	22	44.0%
計	50	100.0%

（出所）筆者の今回の調査による。

は、比べられないものを比べていないか、という点である。目標達成度を数値化した後は必ずそれらを比べること（一覧表を作成する、数字の大きさによってカテゴライズして集計する、など）になるため、フロー指標とストック指標が混在していたり、性格の異なる目標値が混在していたりすれば、無理な比較をしていることになる。

なお、目標達成度を計算しない場合も、目標達成状況を解釈・分析する仕方が一律であれば同じ問題の可能性があるが、本稿では計算する場合のみを検討する。

目標値の性格については、目標達成度の計算において、それを考慮に入れたと考えられる事例は残念ながら全く見られず[8]、ここではフローとストックの区別を取り上げる。

フローとストックの混在が大きな問題となるのは一律に水準達成度の計算を行う場合である。ストック指標の水準達成度の計算では、分子・分母に期首時点ですでにあった水準が含まれており、目標未達の場合達成度が「嵩上げ」された状態になってしまう[9]。

一律に水準達成度の計算を行っている12件[10]は、うち7件がフロー指標とストック指標の区別を全くしておらず、残りの5件は指標をフローとストック、または単年度と累計で分類しているものの、計算する際に区別をしておらず、結局12件のすべてにおいて目標達成度を比較できる状態にない。

一律に変化達成度を計算する場合も、質又は量の同程度の注力が指標値の現状維持をもたらすのが自然なフロー指標の変化と、指標値の一定ペースの増加をもたらすのが自然なストック指標の変化を同列に比較するのが適切でないことは明らかである。ところが一律に変化達成度の計算を行う12件は何れもフロー指標とストック指標の区別を行っていない。

2種類の計算式を使い分けている3例のうち、宮城県はフロー指標については水準達成度、ストック指標については変化達成度の計算を行うことで、フローとストックの区別を目標達成度の計算にも反映させている。ただし、この2種類の目標達成度は異なる定義すなわち意味をもつものであるから、その数字の解釈においても区別が必要となる。長崎県では、単年度の指標と累計の指標とで宮城県と同様の扱いをしているが、残念ながら多くのストック指標が「累計」指標として扱われていない。

三重県は目標達成度の計算において、最も問題の少ない例であろう。単年度の指標については水準達成度を計算し、累計指標については1年間の変化達成度を計算する。後者は単年度の指標に変換して水準達成度を計算するのと同等である。単年度指標がフロー指標であり、ストック指標を累計指標と同様に扱えば、フローとストックの問題が発生しないが、残念ながら少なからぬストック指標が単年度指標として扱われている。

なお、水準達成度と変化達成度を常に両方計算し、プログラムや指標ごとにそれらを組み合わせて達成度を解釈するのは1つの良い方法だが、今回はそのように併用している事例はなかった[11]。

⑦例外的な場合の目標達成度

ここでは小さい値が望ましく、基準値より低い目標値を設定している場合の水準達成度の計算を取り上げる。そのままでは目標未達の度合いが大きいほど達成度の数値が大きくなってしまうケースである。水準達成度を計算している15事例において採られている方法を表7に掲げた。計算方法が不明の1件を除いてすべてに問題がある。

定義式の分子（実績値）と分母（目標値）を反転させる方法が最も多いが、分数の意味が曖

表7　水準達成度の例外処理

分類	件数	割合
目標値÷実績値を計算	7	46.7%
2－（実績値÷目標値）を計算	3	20.0%
変化達成度を計算	3	20.0%
実績値－基準値を計算	1	6.7%
不明	1	6.7%
計	15	100.0%

（注）不明の1件は調査時点で実績値が公表されておらず、計算方法が確認できないもの。
（出所）筆者の今回の調査による。

昧（実績値が目標値の近傍であればそれらしい値にはなる）で、通常の水準達成度の数値（ベースが異なる比率）との比較も不可能である。

「2－（実績値÷目標値）」を計算する方法も、例えば目標値が100で実績値が110の場合、達成度は90%となりそれらしいが、90%が何の何に対する比なのか意味不明である。

このような場合だけ変化達成度を計算する場合（3件）、その数字の意味は明確だが、他プログラムの達成度と比較できない。「実績値－基準値」（1件）は目標値が関わらない数値である。

筆者が推奨するのは、「1－（実績値÷目標値）」を計算する方法である。上記の数値例では－10%となり、目標値に対して10%分だけ未達であることを意味する。通常の指標（大きい値が望ましく、基準値より高い目標値を設定）については正負を逆転させた「（実績値÷目標値）－1」を計算すれば目標値の何%だけ目標超過・未達かという観点から比較することができる（小野2011）。

⑧外部要因の把握

指標の目標達成状況に関する分析の中で、外部要因について一定の記述や吟味があることが確認できたのは8件（宮城、新潟2件、富山、奈良、愛媛（予算施策評価表）、高知、長崎）である。前回調査では「目標達成に影響を及ぼす外部要因の把握」があるのが6団体としているが、これに相当する。13件の制度・システムでは外部要因等に言及する場所が評価書の中に全くない。

表8　外部要因を把握しているか

分類	件数	割合
記述や分析がある	8	16.0%
何らかの言及がある	29	58.0%
全くなし	13	26.0%
計	50	100.0%

（出所）筆者の今回の調査による。

⑨目標達成度の評定

目標達成状況や進捗状況を何段階かにカテゴライズする評定は全50件のうち38件で行われている。表9は評定に用いられている方法を分類したものである。

表9　目標達成状況の評定

分類		件数	割合
達成度計算あり	達成度の率による階級分け	18	36.0%
	一律の水準達成度	5	10.0%
	一律の変化達成度	10	20.0%
	水準達成度と変化達成度の使い分け	3	6.0%
	達成・上昇・下降などの評定	2	4.0%
	達成度＋α で評定	1	2.0%
	複数指標の目標達成割合で評定	1	2.0%
	複数指標の達成度＋α で評定	2	4.0%
	その他	1	2.0%
	評定なし	2	4.0%
計算なし	何段階かに評定	5	10.0%
	達成／未達の分類	3	6.0%
	複数指標の目標達成割合で評定	1	2.0%
	複数指標で何段階かに評定	4	8.0%
	評定なし	10	20.0%
計		50	100.0%

（出所）筆者の今回の調査による。

評定の目的は、指標及びプログラムの目標達成度や進捗状況を端的に示すことであり、個々の指標やプログラムの進行管理の観点からも、統一的で明確なルールが要請される。そのことはまた指標間やプログラム間で目標達成状況を比べることを可能にする。評定に基づいて総合計画など全体の状況を把握するためには、この比較可能性が前提条件となる（この点は次の項目⑩で述べる）。

目標達成度の計算を行う27例では、25件で評定がなされ、指標ごとに評定するのが21件、複数の指標を含む施策などの単位で評定するのが3件である。最も多いやり方は目標の達成度の率による指標ごとの階級分けである。このような評定の結果は（数字を割り当てれば）順序尺度となる。評定結果に間隔尺度の性格をもたせたい場合（典型的には平均を計算する場合。A・B・Cなどの記号で表す場合はその後の取扱い次第である。）は、等間隔の階級にするのが望ましいだろう。

このように評定のルールは明確であっても、肝腎の目標達成度の率が比較可能でなければ、

評定結果も比較可能にならない。フロー指標とストック指標の混在は（厳密に言えば、達成度計算の妥当でない例外処理や目標値の性格の混在なども）比較可能性を損なう。達成度の率による階級分けをしている18件のうち、一律に水準達成度を計算している5件は項目⑥において述べた理由により深刻な問題がある。一律に変化達成度を計算している10件も、深刻さの度合いは低いが問題がある。計算式を使い分けている3件も、残念ながらそれらを区別することなく階級分けしてしまっている。

達成か否か、未達の場合は上昇・下降の何れかなどによって分類する場合（2件）は、達成度の率の比較可能性が問題にならないが、計算した率そのものを活かしていない。

達成度の率から直接階級分けするのでなく、定性的な要因等も加味して評定する（1件）のは比較可能性の問題を回避できる1つの方法である。しかし、計算した率をもとに如何なる情報を加味して評定したのかを明示しなければ本当の解決とはならない。

複数の指標を含む施策などの単位で評定する3例のうち、目標達成した指標の割合による評定（1例）は、達成度の率を活用していないものの、意味が明確ではある。達成度の率に他の要因を加味した評定（2例）では、上と同様、評定の根拠の説明が重要となる。

目標達成度の計算を行わない23例では評定を行わないものが10件ある。この場合、目標達成状況の簡潔な表示が全くない分、より本格的な解釈・分析が求められる。指標ごとに目標達成・未達で分類する（3件）、施策単位で指標の目標達成割合によって評定する（1件）のはともに意味は明確だが、目標値にどれだけ近いかといった達成度合いの情報がない。指標単位・施策単位で定性要因も加えて評定する場合（9件）の成否は評定の根拠次第であろう。

総じて、評定結果の品質は高くないと言わざるを得ない。事態は改善が必要である[12]。

⑩目標達成度の集計・集約

総合計画の全体など、大きな括りでの目標達成状況の集約・集計は37件で行われている。表10は集計・集約の方法を分類したもので、指標の達成度の率の平均を直接計算しているのが2件、指標単位の評定結果を集計・集約しているのが26件、施策等の単位での評定結果を集計・集約しているのが8件などとなっている[13]。

また集計・集約の形態をみると、何らかの平均値を計算するものが7件、評定結果をカウントして集計結果を表やグラフにするものが21件、評定結果を指標や施策等の一覧表の形で集約するものが8件などとなっている[14]。

目標の達成度の率から平均値を直接計算する方法（2件）は、達成度の率が比較可能であれば最善の策であるが、1件はフロー指標・ストック指標が混在した水準達成度であり、もう1件は目標達成度の2種の計算式を区別せずに平均の計算をしており、比べられないものを比べる形となっている。

評定結果に数字を割り当て平均を計算する方法（5件）は、評定結果の比較可能性のほかにも、その数字が少なくとも間隔尺度とみなせることが明確な意味を持つ条件となる。ただし、達成度の率が互いに比較可能な形で計算されているのであれば、比例尺度の率をわざわざ間隔尺度の評点に加工したりせずに直接平均を計算する方がよい[15]。ちなみに、この5件は何れもフロー指標とストック指標が区別なく混在し（4件は一

表10　目標達成状況の集計・集約

分類		件数	割合
指標単位	達成度の率→平均を計算	2	4.0%
	率による階級分け→点数化→平均計算	5	10.0%
	率による階級分け→カウント	8	16.0%
	率による階級分け→一覧表	3	6.0%
	達成／未達等による評定→カウント	3	6.0%
	達成／未達等による評定→一覧表	2	4.0%
	達成度の率によらない評定→カウント	4	8.0%
	達成度の率によらない評定→一覧表	1	2.0%
施策等単位	施策等の評定→カウント	6	12.0%
	施策等の評定→一覧表	2	4.0%
その他		1	2.0%
集計・集約なし		13	26.0%
計		50	100.0%

（出所）筆者の今回の調査による。

律に水準達成度を計算、1件が一律に変化達成度を計算)、評点が明らかに間隔尺度とは見なせないものも3件あった。

評定結果をカウントして集計する方法、一覧表の形で集約する方法（これも比較の一形式である）の成否は、比較対象である評定結果の比較可能性次第である。項目⑨で述べたことから明らかなように、これらの集計・集約結果の品質には多くを期待できない。

なお、評定結果を集計したりせずに一覧表の形で提示することにとどめれば、比較の困難な評点を数えた結果という不確かなデータを提示しないという利点があるとも言えるが、それは達成度の評定の品質が不十分であると自ら認めたともいえる。

4. 業績測定型評価の今後に向けて

今回の点検の結果、個々の指標の目標管理においても、目標達成状況の指標間の比較や計画全体など総括的な集計・集約においても様々な課題が明らかになった。前回の点検からいくらか改善の見られる項目もあったが、全体として芳しくない状況にあると言わざるを得ない。

業績測定型評価によってPDCAサイクルを回す、その中核的機能である数値の指標による目標管理について、本稿が指摘した問題群を改善することなく、本来の意味で徹底することはできないはずである。品質の良くない指標や目標値、目標達成度の集計・集約などに基づくPDCAサイクルが良い結果をもたらすことは期待できない。

以下では、状況の改善に向けとりわけ重要と考えられる4点を指摘したい。

(1) 指標の妥当性の向上

今回の調査では、指標の妥当性に係る情報が示されているか否か、主として評価書や計画書のフォーマットやルール、記述状況という観点から点検しており、設定された指標自体の妥当性を直接吟味することはしていない。しかし、大半の評価書や計画書において指標の妥当性に関する記述が不十分または全くないということは、少なからぬ指標の妥当性自体が危ういと考えるべきだろう。

社会科学において測定の妥当性を巡る議論は様々あるが、業績測定型評価における妥当性の第一はロジックモデル上の位置づけが明確かつ適切であることであろう（Hatry1999、2006)。その上で明確かつ適切な指標定義が求められる。しかし、自治体の評価作業において指標設定のためにロジックモデルを実際にきちんと描くことなどを求めるのは（ごく少数の特筆すべき実例はあるとしても）ハードルが高いだろう。そこで提案したいのは指標設定に関するチェックリストの活用である。

欧米で幅広く参照される業績測定型評価のテキストやマニュアル類ではしばしば妥当性（ただし妥当性が含む概念は様々である）に信頼性など他の条件を加えて評価指標が満たすべき条件が掲げられている（例えばHatry1999、2006、Ammons Ed.1995、ASPA2000)。また田中啓はHatry及びAmmonsによる整理を統合した基準を提示している（田中2014)。これらは、ほぼそのまま評価指標を設定する際のチェックリストと考えることができる。表11は自治体の評価作業において使うことを念頭に、簡便さを追求したチェックリストとして筆者が提案したものである（小野2015)[16]。

表11　指標の妥当性のチェックリスト

基準	説明
①目標との適合性	ロジックモデル上の位置が明確・的確か
②包括性	政策の重要な側面をカバーしているか
③外部要因の影響	他の政策、他機関の政策、行政の外の要因などの影響がどれくらいか
④感度、感応度	本来の効果（結果）の大小に指標値が反応するか
⑤計測可能性	実際に安定的に測定できるか
⑥適時性	タイムリーに測定できるか
⑦費用	データ収集コストが過大でなく、予算内
⑧悪影響の可能性	局所への集中、データ操作などの可能性

（出所）小野（2015）

なお、表11において「妥当性」とは本稿において「測定すべきものを測定していること」と定義した通りの広義のものであり、上記の文献では信頼性など別の条件として分類されているものを含む。

(2) 指標間の比較可能性の向上

目標達成状況を総合計画全体や政策分野ごとに集計・集約する場合はもちろん、個々の指標や事務事業の目標達成度を把握する場合にも、指標間の比較可能性がきわめて重要な意味をもつことはすでに述べた通りである。そのための工夫をしている団体もあったが、十分な比較可能性を確保している例は残念ながらなかった。

この問題の解決策としては、以下の通り数字の解釈において比較可能性を確保する方法Aと、数字自体を比較可能なものとする方法Bの2通りが考えられるだろう。方法Bはさらに達成度を測る目標の類型によって3通りに分類できる。

A. 2種類の目標達成度を計算

すべての指標について、水準達成度・変化達成度の2種類の目標達成度を計算し、目標値の性格やフロー・ストックの別などを踏まえ、2つの達成度を組み合わせて解釈・分析する。それを踏まえて評定することにより、比較可能な評点が得られるだろう。

B-1. ストックをフローに変換し単年度を評価

年度ごとの目標値がある場合、ストック指標の年度ごとの差分（年度の期首・期末の差）を取ることによってフロー指標に転換すれば、すべての指標の目標達成度を、1年間の実績に関する評価という観点から直接比較できる。各年度の目標達成度は、水準達成度を計算することで得られる。

B-2. ストックをフローに変換し累計を評価

上と同様にストック指標をフロー指標に変換する。計画など一定期間の実績に着目する場合、各年度の実績値（ストック指標については変換後の実績値）を対象期間について合計し、水準達成度を計算する。もとがストック指標の場合は変化達成度と同じ計算をしていることになる。すなわちフローの累計とストックの変化の比較である。

B-3. ストックをフローに変換し変化を評価

上と同様に、ストック指標をフロー指標に変換する。計画など一定の期間において、どれだけ質・量のパフォーマンスを向上・改善できたかを問うとすれば、期首から期末までの（変換後の）指標値の変化に関する達成度すなわち変化達成度を計算すればよい。

なお、B-1～ B-3については、総合計画等の進行管理において、組み合わせて用いることにより指標やプログラムの多面的な目標管理が可能になるであろう。

(3) 補助情報の重要性の認識と開示の徹底

一般に数量データを取り扱う際の基本原則の1つ、数字と補助情報は両者の組み合わせで初めて数量データとしての意味を持つということは、業績測定型評価においても常に強調すべきことであろう。指標や目標値の選択・設定の根拠や目標達成度の計算や評定の方法の詳細など、これは本稿で論じた10の点検項目すべてにも関わるが、それら補助情報が不十分であったり、全くなかったり、その現状はすでに指摘した通りである。補助情報の明示が必須のものとなれば、目標管理の品質に関する内外のチェックが機能しやすくなり、品質の向上を促すことに繋がるのは確実である。

(4) 立案段階における目標設定

本稿では目標値の設定について専ら評価指標の観点から論じたが（点検項目③④⑤）、問題点が多い現状は、プログラムの立案段階における課題を示唆する。プログラムの立案時にそのニーズや目指すべき水準を明確に把握していれば、その時点で妥当性を有する評価指標を設定し、同時に設定根拠や性格が明らかな目標値を設定することが可能になる場合が多いはずである。

おわりに

業績測定型評価が機能するためには、当然その基本ツールである評価指標やその目標値の品質が一定程度以上のものでなければならない。

しかし都道府県を対象とした本研究の結果を見れば、市町村においてもこの条件を満たす団体は多くないと推定できる。

指標と目標値の品質が良いからといって業績測定型評価が機能するとは限らないが、その懸念は無用かもしれない。むしろ、評価を積極的に遂行する意思の欠如に起因して指標・目標値の品質が改善しないままという姿がときに垣間見える事態こそが深刻である。

なお、本研究における指摘や提言は一義的には内部評価の実施主体たる行政に向けたものであるが、行政の外から二次評価や外部評価を行う場面の関係者にも参照していただきたい。外部から評価指標・目標値の問題点を指摘することは、個々のプログラムを巡る議論に止まらず、評価制度・システムの改善を促す良き圧力になるはずだからである。

謝辞

本論文の初稿に対して匿名の査読者から大変有益な指摘をいただいた。深く感謝の意を表する。

注記

1 本稿では、公共部門の政策・施策・事務事業（広義政策）や様々な活動・サービス等を評価する制度・システムを「行政評価」と総称する。

2 同調査によれば、かつて実施していた行政評価を廃止した団体は52（政令市1、市区27、町村24）ある。

3 業績測定を狭く定義するとすればそれは指標のモニタリングであるが（ハリー・ハトリーは広く読まれているテキスト（Hatry1999、2006）の冒頭で業績測定を「サービスやプログラムの結果と効率を定期的に測定すること」と定義している）、本稿では、目標値の設定から目標達成度の把握と分析、対応策・改善策の検討、それらの情報の公表に至る一連の過程を含めて業績測定型評価と呼ぶ。

4 本稿で論ずる評価指標は数値の指標である。業績測定で用いる指標は数値の指標であることが当然の前提である一方、現実の業績測定型評価において「定性的」な指標が一定程度用いられていることもまた事実であるが、本稿では取り上げない。自治体での定性指標の使用は比較的少ないが、国の政策評価においては定量指標と定性指標の役割分担も重要な課題となっている。

5 「経済財政運営と改革の基本方針2015」（2015年6月30日閣議決定）では、行政に対する定量的な評価に係る取組が十分でない、エビデンスに基づくPDCAの徹底をせよ、と書かれている。参議院は2015年7月の本会議で「政策評価制度に関する決議」を可決し、目標管理型政策評価についても測定指標の更なる開発・設定などを求めた。また会計検査院による2015年12月の政策評価の実施状況に関する報告では、測定指標・目標値設定の改善や行政事業レビューとの更なる連携が求められた。一方、総務省でも2015年5月に政策評価審議会のワーキング・グループ（筆者もその一員）が設置され、目標管理型の政策評価に係る改善方策の検討が2016年度にかけて進行中である。

6 本研究の主たる目的は、業績測定の制度・システム設計を論ずることにある。また現実問題として、1つの団体につき数十から数百程度設定されている評価指標そのものの点検は、本研究の範囲では不可能である。

7 「水準達成度」・「変化達成度」は、前回点検の報告（小野2008）及び業績測定型評価に関する拙稿（小野2011、2013）においてそれぞれ「期末型達成度」・「期間型達成度」と呼んでいたものである。

8 例えば、非常に高く挑戦的な水準の目標値を掲げている場合は、水準達成度よりも変化達成度によって進捗状況を測るべきであろう。一方、達成が必須の目標値であれば、どれだけ近づいたかではなく最終的にどうなったのかを水準達成度によって測るべきであろう。

9 例えば、ある自治体の基準時点の人口が9万5千人、その時点で人口増加を目指して設定した目標値が10万人とする。目標年度の実績が9万人に減少した場合、人口増加策の水準達成度は90%となる。

10 愛媛県の事務事業評価表では、過去3年程度の実績値に回帰直線をあてはめ、評価対象年度の推計値に対する実績値の比を達成率として計算している。その計算式から本稿では水準達成度と分類している。

11 前回の調査では岡山県が、すべての指標について「達成率」「進捗率」という2通りの達成度を計算し

ていた。この「達成率」とはフロー指標は水準達成度、ストック指標は1年間の変化達成度（フロー指標に変換した水準達成度に等しい）であり、「進捗率」は変化達成度であった。

12 前回調査では項目⑨（評定）と⑩（集計）を合わせて集計しており、本稿では言及しない。

13 1つの評価制度・システムにおいて、指標単位の集計・集約と複数の指標を含む施策単位などの集計・集約の両方がある場合は、指標単位の集計・集約のみを取り上げた。

14 1つの評価制度・システムにおいて、一覧表、カウント、平均計算のうち複数がある場合、加工度の高い順（平均計算、カウント、一覧表の順）に1つを取り上げた。

15 達成度の率を階級に分けて評定することは、連続量のデータの情報を分かりやすさと引き換えに捨てていることになる。それら評点の平均値を計算するのは、捨てた情報を使わずに再び連続量を作ることになる。例えばオリンピックの陸上競技の着順のように、順序尺度の数字が決定的に重要な場合（オリンピックの陸上競技であれば記録よりもメダルの色のほうが重要であろう）以外、そのような平均値は奇妙なものである。

16 筆者も参画して開発した大学評価指標の設定のためのチェックリスト（独立行政法人大学評価・学位授与機構2015、渋井他2016近刊）の検討過程における議論も参考にしている。

参考文献

小野達也（2008）「都道府県の評価の10年－定量評価のこれまでとこれから」『日本評価研究』8(1)

小野達也（2011）「業績測定型評価における目標設定と達成度評価の妥当性－行政評価の形骸化を避けるための条件」『地域学論集』（鳥取大学地域学部紀要）8(2)

小野達也（2013）「政策評価と実績測定－府省の実績測定における計量・計数の現状」『日本評価研究』13(2)

小野達也（2015）「客観的かつ具体的な政策効果の把握及びそのための適切な目標・測定指標の設定について」総務省行政評価局『平成27年度政策評価に関する統一研修（地方研修）テキスト』

佐藤徹（2012）「自治体評価における『おかしな成果指標』の実態分析」『日本評価学会第13回全国大会報告資料』

渋井進・浅野茂・橋本貴充・小野宏・小野達也・山﨑その・田中弥生（2016近刊）「自己評価力向上支援のための評価指標設定に関するチェックリストの開発」『大学評価・学位研究』第18号

総務省（2014）『地方公共団体における行政評価の取組状況等に関する調査結果』
<http://www.soumu.go.jp/main_content/000278815.pdf>
（2016年1月1日アクセス）

田中啓（2014）『自治体評価の戦略 ―有効に機能させるための16の原則』、東洋経済新報社

独立行政法人 大学評価・学位授与機構（2015）「指標のチェックリストと使い方」EAワークショップ『指標の選び方＆指標信頼性・妥当性のチェックリスト』資料5
<http://www.niad.ac.jp/n_kenkyukai/1259551_1207.html>
（2016年1月1日アクセス）

American Society for Public Administration (2000). *Performance Measurement: Concepts and techniques*. Washington. D.C.: ASPA Center for Accountability and Performance.

Ammons,D.N. Ed. (1995). *Accountability for Performance: Measurement and Monitoring in Local Government*. Washington. D.C.: Internation City/Count Management Association.

Hatry,H.P.(1999). *Performance Measurement: Getting Results*. Washington, D.C.: Urban Institute.（上野・上野訳（2004）『政策評価入門－結果重視の業績測定』、東洋経済新報社）

Hatry, H. P.（2006）. *Performance Measurement: Getting Results 2nd Edition*. Washington, D.C.: Urban Institute.

（2016.6.23 受理）

Current State and Future Agenda for Performance Measurement in the Local Government: Lessons Learned from 20 years Practices in the 47 Prefectures

Tatsuya Ono

Faculty of Regional Studies, University of Tottori
onotatsu@rs.tottori-u.ac.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 performance measurement to implement Plan-Do-Check-Act cycle in the management. However, there seems to be a number of problems in their way of handling data in the process of setting performance measures and their targets, and benchmarking. This article shows the results from the survey of performance measurement systems in all the prefectures of Japan, which follows the previous survey in 2008. This article classifies the typical problems of performance measurement and also explores the way of acquiring the quality of measurement and benchmarking.

Keywords

performance measurement, prefectures of Japan, performance measures,
target setting and benchmarking

【総説】

自治体評価を振り返る
－「活かさず殺さず」の20年－

西出　順郎
岩手県立大学
nishide@iwate-pu.ac.jp

要　約

本稿は、自治体評価の20年の歴史を研究及び実務の視点から振り返り、今後の自治体評価について若干の検討を加えるものである。はじめに我が国の自治体評価の特徴である組織管理への接近について、その要因を自治体評価が台頭した2000年前後の自治体を取り巻く環境から抽出する。次にそれゆえの自治体評価の課題、課題克服のための処方箋、さらには処方箋の限界等を記述する。最後に自治体評価の近年の動向を考察し、自治体評価の行方について、その方向性を模索する。

キーワード

地方自治体、業績評価、組織管理

はじめに

1996年に三重県が「事務事業評価システム」を導入し、それ以降、全国の地方公共団体（以下、「自治体」という）の中で、政策等の業績を評価するシステム（以下、「自治体評価」という）が急速に広がった。2013年10月1日現在、計1,060団体、59.3％の自治体が自治体評価を導入している（総務省2014）。この1996年を「自治体評価元年」とするなら早20年の歳月が流れたことになる。

公的部門の評価理論（Evaluation Theory）では、事業の改善や社会への説明責任が主たる評価目的とされ[1]、各自治体もまたその両者を評価目的に掲げている。しかし実際のアンケート調査、例えば総務省（2014）に鑑みると、「行政運営の効率化」、「行政活動の成果向上」、「PDCAサイクルの確立」、「職員の意識改革」など、自治体評価の関心は住民への情報発信よりも組織内部での情報活用に集まっている。「職員の意識改革やアカウンタビリティ以上に、自治体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にある」（小野2008：33）との指摘があるように、自らの管理機能を支援するシステムとして自治体評価を位置付ける論者は少なくない[2]。

導入直後の自治体評価は、無駄な事業を削減する伝家の宝刀としてマスコミからの称賛を受けた。また政策や施策、事業の効果（以下、「政策効果」という）の評価など「できるわけがない」という、職員の固定観念をいとも簡単に覆し、「アウトカム（成果）」という聞き慣れない言葉で行政改革の一翼を担った。「仕事とは、何を『したのか』ではなく、何を『達成したのか』である」。当時はこのような行動規範がNPM（New

日本評価学会『日本評価研究』第16巻第1号、2016年、pp.17-30

Public Management）というムーブメントとともに席巻した時代であり[3]、最も早く取り入れた三重県の評価システムは自治体評価のディファクト・スタンダードとして全国に波及した。

しかし歳月が流れると、磨きをかけない宝刀は錆びつきを始める。評価システムの形骸化や機能不全といった批判が研究者のみならず実務家からも噴出するようになった。また、それらに応える有効な改善策は何ら打たれることもなく、2000年代の後半には「事業仕分け」といった劇場型のコスト・カッターが全国を席巻するに至った。自治体評価は、機能的に活用されず、制度的に廃止されず、結局、組織の中で「活かさず殺さず」のまま、未だ燻り続けている。

自らが統制する組織「内部」の支援機能に過ぎないにもかかわらず、なぜ自治体評価は役に立たないのか。本稿ではこの素朴な問いを改めて見つめ直すことを念頭に研究のみならず実務の視点も交え、特に自治体の運営管理（以下、「組織管理」という）での活用という観点から「自治体評価の20年」を考えてみたい。

1　自治体評価と組織管理

自治体評価はなぜ組織管理の支援機能を期待されたのか。体系的技法を活用し、政策改善の本質を抉らんとする評価の科学的志向に依拠すれば、その期待に偏りがあるのは否定できない。しかし自治体評価が全国普及し始めた1990年後半に自治体が直面した問題に鑑みるとその流れは決して不自然なことではなかった。

当時の「問題」とはどういうものだったのか。既に多くの論者が整理しているが[4]、自らの実務経験を踏まえ筆者なりに振り返ると、次のとおりである。

その第一は国・地方全体に関わる政治的な動きである。バブル経済崩壊による巨額の財政赤字や改革派首長等による地方分権推進の取組は効率・効果的な政策立案を自治体経営に要請した。また財政の効率化については国もその具体的方策を自治体に強く求めた（例えば「地方自治・新時代に対応した地方公共団体の行政改革推進のための指針の策定について」（平成9年11月14日自治整第23号）[5]）。

第二は自治体内部の不正経理の問題である。官々接待等の裏資金を捻出していた「カラ出張」の存在が多くの自治体で露見された。この組織的な不正出張は、例えば福岡県で約59億円、山梨県で44億円、北海道で約42億円と[6]、各自治体の不正流用額が数十億円にも達する大規模スキャンダルへと発展した。地域住民の信頼を失った自治体は自らの襟を厳しく正すとともに質の高い政策の実現による信頼回復に迫られた。

第三は政策を統制する伝統的な仕組みの金属疲労である。これは国・自治体の財政赤字に起因するものであったが、この状態を何とかしようとする動きは、三割自治が示唆するように受身的な仕事が多い自治体にとって、数少ない自律的な変革への試みであった。このことに自治体評価の観点から着目することはあまりなかったので、少し丁寧に説明しておきたい。

自治体の税収入が右肩上がりの時代には、各事業部門が前年度を上回る予算要求をしてきても幾ばくかの財源の上乗せは可能であった。その差配を駆使して財政部門は自治体の政策全体を統制し、事業部門も程度の差はあれ毎年度の予算の上乗せで納得することができた。要はこのさじ加減が財政部門の組織内権力の源泉だったのである。但し、当該権力を駆使、維持するには組織内の政治的な均衡が重要であり、決して政策効果の客観性や科学性が厳しく追求されることはなかった。時には事業部門の総務係長といった予算担当者の政治的交渉力に差配結果が大きく依存した。

この予算の差配プロセスは御恩と奉公を彷彿させるパターナルな関係性を財政と事業部門との間で構築させ、確かに財政部門は組織の政策統制力を温存し続けることができた。また当時の企画部門は総合計画の集約・加工といった業務が中心であり、当該部門による予算査定や編成への介入は皆無に等しかった。しかし、政策的に必要と信じるが満額の予算を獲得できかった事業担当者もしくは無理難題な宿題のこなし具合を見極めながら予算要求の本気度を確かめる財政部門の予算査定の仕振りを嫌う職員は、

この政策統制のありかたを是としてこなかった。

長期にわたり財政事情が好転せず予算削減が余儀なくされる状況が続くと、差配による政策統制は従来のように機能しなくなってきた。財政部門の姿勢が予算の上乗せといった便益の差配どころか、常に前年度の予算を削減する立場へと転じたからである。そこで事業部門は上乗せなきパターナルな関係維持に反発し新たな行動規範による政策統制の仕組みの必要性を主張しはじめ[7]、同時に財政部門は客観性や科学性を理由にできる予算削減の仕組みを模索しはじめた[8]。

すなわちこれらの動きがいわゆる「政策の風」として自治体内を席巻し、新たな仕組みの原動力となった自治体評価はマスコミ等[9]の後押しを受けながら劇的に表舞台へと駆け上っていったのである。ゆえに「政策効果の数値化」、「政策効果による政策形成」という新たな行動規範が自治体内部で強烈に支持され、自治体評価が政策や事業の設計や予算の査定・編成の支援システムとして全国に浸透していったのは、自然の成り行きだったといえる。

2　役に立たない理由

自治体評価は自律的に導入され20年以上の歳月を経ているにもかかわらず、なぜ未だ抜本的な改革が進まないのか。その答えは決して難しくない。一度でも評価の実務を経験すれば、もしくは評価システムの仕組みを精査すれば、誰もが気づくであろう。なぜなら評価作業を完遂させるシステムとしての機能設計がほとんどブラックホールとなっているからである。その具体をみてみよう。

第一に、評価の作業ルールはあっても評価結果をどのように反映させるのかといった使い道の基本ルールが存在しないこと、さらにそのようなルールは実際に作りようがないことである。万人が合意する便益計算ルールのもと、全事業の便益が事前かつ工学的に産出できれば可能かもしれないが、現実的にはあり得ない。業績測定の結果ですら、目標値の達成イコール事業費の倍増となるのか未達成イコール事業費倍増となるのか、測定結果の使い道を簡単にはルール化できない。すなわち1960年代の米国PPBS（Planning Programming Budgeting System）の失敗が示すように、「選好の自動化」といった予算編成は決して単純なことではないのである。その意味では、予算要求の本気度を試す財政部門と事業部門との長時間にわたる予算折衝はあるべくしてできた仕組みともいえるし、事業部門が評価のシステムを以って非効率な予算折衝を何とか合理化させたいとするロジックも十分理解できよう。

勿論、事業の有効性を議論するツールとして何らかの活用ができれば、予算編成等に何らかの貢献はできるかもしれない。しかしそれすら簡単なことではない。なぜなら事業の有効性を測る評価基準の設定すら複雑化しているからである。これは第二としての公的サービス特有の問題であり、評価の物差しなど容易に特定できないことを意味する。例えば、しばしば引き合いに出される公立図書館の評価では「貸出冊数」や「来館者数」（それらの数が多いに超したことはないが）といった活用指標を単純に主たる目標とするわけにはいかない。近年不登校者の「駆け込み寺」として脚光を浴びるように、図書館の便益は多様だからである。高価な学術書を一般市民が容易に閲覧できることも図書館の特長である。とはいえ不登校者の利用者数や高額学術書の利用者数を目標指標とすることには誰もが違和感を覚えよう。このように教育や福祉など公的サービスの多くは目標の単一化や数値化にうまく馴染まないのである。

仮に数値指標が特定できても十分とはいえない。目標値の持つ政策的な重み及びその責任の所在の曖昧性、これが第三の問題である。通常、数千にも及ぶ事業の目標値は全て事業部門の主導のもと設定される。それゆえに着実に実現可能な目標値が選択される。達成不可能な目標値は目標としてそぐわないし、背伸びすれば届くかもしれない目標値であっても選択には二の足を踏む。そもそも地域住民が具体に望んだ数値などないし、無謬性の理屈に立てば未達成イコール行政自らの失敗と考えられる。選挙公約で掲げた数値目標がそのまま組織の目標へと転じ

ていれば別だが、個々の事業の目標値にまで首長が積極的に関心を持つことは少ない。その空気は職員も十分嗅ぎ取っている。よって事業部門自らが目標値を選択する自治体評価のシステムにおいてはエージェンシー理論でいう「本人・代理人関係」が不明瞭となり、目標管理のための緊張感やインセンティブが創出され得ないのである。

第四としては自治体に内在する統治構造上の制約が挙げられる。自治体の事業には法令に基づく事業や補助金等による国の政策誘導的な性質を持つ取組が多い。国の法令を超える独自の規制方策も展開できない。すなわち事業の質・量的管理の統制に関しては自ずと限界があり、目標値は国との政策上の関係性を斟酌し、限られた裁量の範囲内で設定しなければならないのである。また、仮に自治体独自の一般事業だからといって容易にできるわけでもない。自治体の財政力もしくは事業規模には一定の限界があるからである。例えば、毎年50万円の補助金が地元商店街の活性化に本当に寄与するのであろうか。政策アイディアが豊かであればともかく、実際にこの予算で商店街の活性化を導くことは想像し難い。地域の社会的変容を発現させるほどの事業予算を準備できないのである。自主事業であるにもかかわらず事業と政策効果の間に存在する因果的連鎖（ロジックモデル）を説明できないのは、政策効果と事業規模の関係性に起因することが多く、これが目標値の選択をより保守的にさせる一因ともいえる。

最後はレント・シーキングの問題である。事業部門としては、政策効果がなくても政治的に廃止できない事業、さらには自らの便益追求の都合上、政策効果が曖昧でも存続させたい事業が存在しないわけでもない。上記の商店街への補助金事業などは過去から延々と続く事業で、廃止すること自体が政治問題化する懸念があるがゆえに継続しているのかもしれない。第5節で示すが、事業部門が評価の主たる従事者である以上、事業部門は、予算を増大化するもしくは維持する行動の一環として、時には評価結果を故意に歪め評価結果の客観性を自ら放棄してしまうのである。

以上、いずれの問題も評価システムのアウトカムを達成するうえでの懸案でありながらほぼ等閑視され、さらにはなんら有効な処方箋も抉り出せず、それらの解決や調整が現場の事業部門に丸投げされていると考えるのが、現行システムの現実といえよう。

3　改革・改革への処方箋

自治体評価の改革について、この20年間何ら無策であったというわけではなかった。近年はそれらの動きすら鈍化してきたが、全国的に本格導入されるとともに多くの研究者や実務家が改善策を模索してきたのは事実である。どのような策が提示されたのか。以下、それらを整理してみよう。

最も基本的な処方箋といえるのが「自治体職員の評価能力の向上」である。小野（2001）、白川・小野（2002）、田中（2003）、西出（2003）、田渕（2005）などは、各職員の評価能力が向上することで適切な指標の設定や評価結果の質の向上を期待した。なかでも自治体評価の主目的が組織管理の支援機能にあるため、評価の技術的スキルのみならず、組織内における評価情報の有効活用を支援するマネジリアル・スキルが評価の力量の一つとして強調された（西出2003）。

二つめは「評価の汎用性拡大」である。これは個々の事業を細かく精査する事務事業評価への批判であり、プランニング機能への支援強化を目論むものであった。例えば佐藤（2001）は、評価単位が小さく数も膨大であるゆえ事務事業評価は行政資源の最適配分や事業再編等の政策判断に使えないとし、施策レベルへの評価の拡大を主張した。プランニングと評価の連携については米国のGPRA（Government Performance and Results Act）に導入された戦略計画（Strategic Plan）や計画管理ツールとしての評価の活用が着目された。例えば上山（2001）は計画策定やミッション志向の評価の重要性を論じ、また北大路（2001）は、評価の原点は組織ミッションの達成度の評価にあるとし、いずれも実施責任者が評価の責任を取る仕組みや組織評価への転換

等を指摘した。加えて民間企業で評価された管理手法の導入可能性も議論された。例えば南（2003）は事務事業評価に間接経費を含めたコスト分析機能を持たせるためABC（Activity Based Accounting）を、大住（2005）は事業の妥当性や実施面での効率性のみならず、SWOT分析を用いた組織マネジメントの取組を提案した。また森田・中嶋（2004）はマネジメントシステムの具体としてBSC（Balanced Scorecard）を、大住（2002）や田中（2005）は自体体間の事業比較を目的としたベンチマーキングシステムの導入可能性について論じた。

三つめは「実績評価からの解放」である。例えば西出（2005）は実績偏重の評価の基準を批判し、事業の持つプログラムセオリーの事前検証によって予算査定等での有効活用を主張した。また山谷（2007）は我が国の業績測定の問題を踏まえ、ユーザーの様々なニーズに応える応答型の評価やセオリー評価の活用を、同様に窪田（2008）は事務事業評価を自己評価による安価かつ網羅的に体系的に導入しようとした試みとしセオリー評価やプロセス評価への転換を提言した。

最後は「組織学習としての評価の重視」である。本来は小野（2008：23）が指摘するように「評価作業自体を通じて意識改革をなされるといった効果もあろうが、それはあくまで副次的なもの」とするのが基本であろう。普及当初の趣旨に鑑みれば、自治体評価が組織学習へとその主目的を移行させることは評価目的としての撤退を意味する。しかしながら自治体評価のディファクト・スタンダードたる三重県の事務事業評価システムが職員意識の改革を意図していたのも事実であった[10]。したがって現行の自治体評価が職員意識の改善を重視するとなっても不自然なこととはいえず、実際に多くの研究者が意識改革のツールとして高く評価した。例えば、高寄（2000）は官僚制を打破する仕組みとして、今井（2001）は管理ツールとしての限界を認めながらも知恵を共有するための手段として、そして真山（2001）は目的と手段の関係を論理的に考える仕組みが政策形成力の向上に資するという観点から、組織学習や意識改革に対する自治体評価の有用性を論じた。

4　処方箋の限界

自治体評価を導入する自治体は着実に増えている。しかしながら抜本的な改革の動きについては現場からは聞こえてこない。その方策はこれまで議論されてきたにもかかわらず、である。なぜか。評価の現場により焦点を当てるとさらなる二つの問題が垣間見えてくる。

その一つは、これまでの処方箋の未熟性、すなわち具体性が乏しく、現場がそれらを実装化し切れないことである。すべきとされるコンセプトは理解できでも、現実機能のためのハウツウが全く不明瞭でその適合性も未知数なのでは、現行の仕組みを凌駕できないのである。もう一つは現場自身が、実は改革をあまり進めたくないこと、すなわち評価という刃を研ぎ澄ますことで自らが怪我を負うのではないかという不安である。

前者の難しさは、第2節で示した評価の問題、すなわち評価結果の使い道や評価指標設定の難しさ、目標値の責任の所在等々を何ら解決できていない点にある。確かに職員や組織の評価能力を養成すれば、統計解析を駆使した評価は可能になるかもしれない。またSWOT分析やABCを付加することで評価システムはより精緻になるのかもしれない。しかしそれらが政策効果の指標やその目標値さらに責任分担を工学的に明示してくれるわけではない。すなわち評価能力を向上させると如何なる成果指標が現実的に抽出できるのか、どのシステムや技法だと地域住民が納得する目標値を設定できるのか等々、それらの解に対し、どの処方箋も技術的には無頓着なのである。実務からみれば、これまでの処方箋は現実と一定の距離感を保ちながら概念的な解を例示したにすぎず、導入を試みるインセンティブを現場内に発現させるものではなかったといえる。

勿論、研究の結果に頼らず、いわゆる実践的な解を模索するのが実務家の使命であり、実務家としての腕の見せ所でもある。しかし評価に限っては必ずしもそうとは言い切れない。これが後者の、改革を進めたくないとする理由である。「形成・実施・評価」といった政策循環過程のなかで、実務と研究が最も接近し、同じベク

トルを向いて議論をできる場が評価である。それは実務が評価の結果に対して最も客観性や科学性に関心を払うからである。実務は評価のシステムを通して政策判断の正当性を標榜したい。そのために研究の分析作法に基づく信頼度を標榜できるお墨付きが重要となってくる。しかし同時に不安も浮上してくる。その客観性と科学性が自らの政策効果を否定するのであれば、社会からの批判、政策課題の政局化を自らが招き入れてしまうからである。よって実務は研究分析の作法によるお墨付きは欲するものの、赤裸々な評価結果を産出するシステムを敬遠することになる。すなわち評価とは実務家にとって両刃の剣であり、いくら評価結果の質を向上できるとしても、実務自らが評価結果を自由に操作できないシステムを持ち込むことは考えられないのである。その意味では評価の質がどのように批判されても、現行の自治体評価の仕組みは実務にとっては極めて都合のいいシステムといえよう。

5　評価の操作性

地域住民には客観性、科学性のお墨付きを得た評価結果を示したい。しかし自らの評価結果によって余計な問題は起こしたくないから評価結果の操作性は維持したい。ではそもそも評価結果の操作性は評価システム上どのように担保されているのであろうか。全ての自治体に合致するわけではないが、現行システムには概ねそのための二つの仕掛けが装置化されている。

まずは評価技法における仕掛けである。評価システムは業績測定とその結果の原因等を検証する要因分析とを要請している。目標値と実績値のギャップだけでは目標未達以外の情報を何ら見出させないため、政策改善に資するのは専ら要因分析の中身ということになる。通常の評価理論でいえば体系的な社会調査の技法を予定されるが[11]、現行のシステムは要因分析の技法について何らルール的な縛りを持たない。すなわち説得力の乏しい主観的な分析結果であっても受容され、自らに都合のよい政策改善の根拠を提示することが可能となるのである。

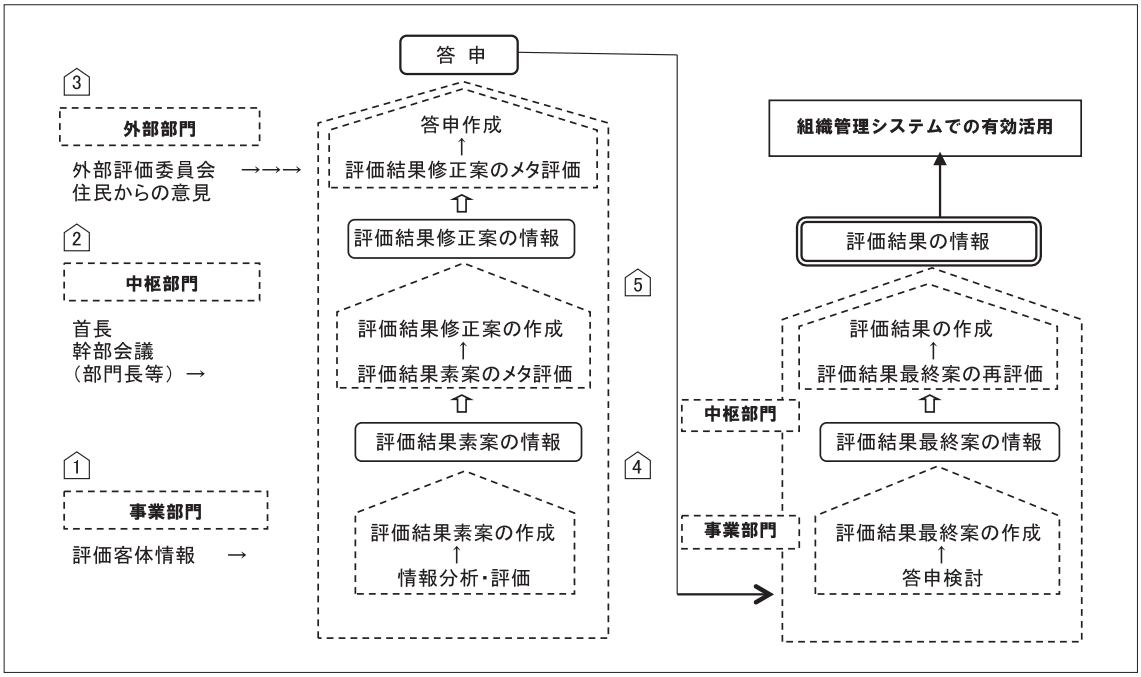
次の仕掛けといえるのが評価結果の決定過程である。一部の自治体では事業仕分けや監査部門型の外部的な評価を実施する場合もあるものの[12]、現行システムでは内部職員が評価結果を作成する。但し評価専担の独立部門が内部に設置されたとしても、組織全体としての評価システムの設計や運用の管理に従事するにすぎず、直接評価を実施することはない。勿論、評価結果の決定・責任主体は条例等のルール上自治体の長であるが、作業としては通常の行政活動と同様のためボトムアップ型の意思決定過程が踏襲される。したがって評価作業に従事するのは、第一義的には評価客体（施策や事業）の実施部門で、評価客体の予算原案を作成・折衝し、予算化された取組を実施する担当者が評価担当者となる。そして彼/彼女らが作業した評価結果の原案を首長が評価結果として了承するまでの過程において、多くの内部職員が複層的に評価結果の仕立て直しに介入することになる。具体的には図1のように、評価客体担当の部門長のみならず副首長をはじめ官房（企画・財政・人事）系及び他の事業部門長で構成される中枢部門、時には外部の評価委員による介入もあり、それぞれが自らの部門の意図に沿った評価結果の形成を目論むのである。

すなわち現行システムは多くの組織内利害関係者の介入を許すことで客観性や科学性とはかけ離れた、各部門の妥協による、もしくは予定調和のとれた評価結果の産出を可能にするのである。最早ここには専門職としての評価者（Evaluator）は存在しない。

6　近年の動向

「自治体評価が役に立たない」。このことは多くの自治体職員から仄聞するし、評価書の質を精査すればほとんどの評価研究者が同意するだろう。計量分析のない評価など、評価ではないとする経済学系の研究者も少なくない。しかしこれらを自治体側が公の場で認めることなどそうあり得ない。すでに導入されている以上、行政活動の無謬性という行動規範の下もし自治体

図1　評価結果への介入過程



（出所）西出（2012）による

自らが「失敗から学ぶ組織」を標榜し評価の機能を正直に否定するものなら、自らの評価すらできない無能力さが批判に晒されるからである。他方、翻って説明責任の立場に転じれば、そもそも評価書の存在自体に、その質の議論は別にしても、意義を見出すことができる。政策の有効性・効率性を省みる機会であることは疑いのない事実だからである。

結局、自治体自身が「自治体評価は役に立たない」と公式に標榜する理由などあり得ず、また使い手たる自治体自身が何らかの役に立っていると主張すれば、研究者や地域住民が評価書の質のいい加減さをいくら主張しても議論の進展は見込めないのである。

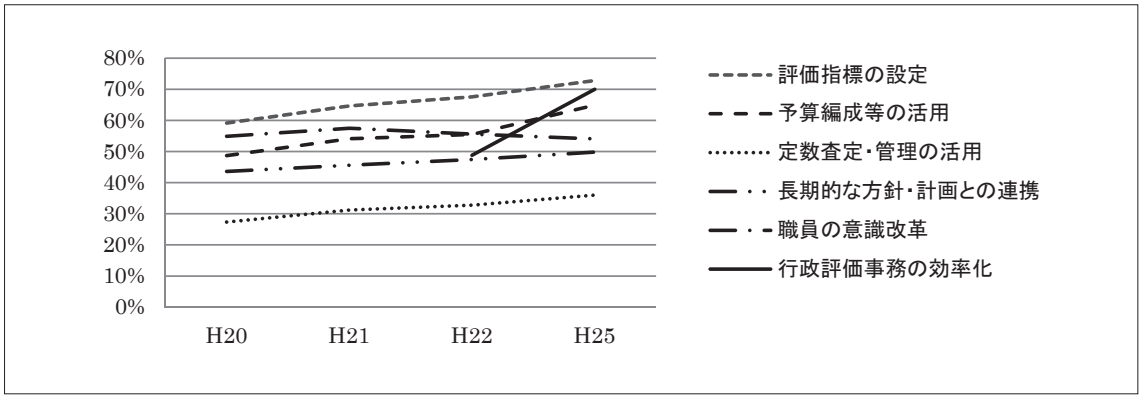
それでも自治体側の発信した情報から「評価が役に立ってないだろう」と想像することは決して難しいことではない。例えば何らかのアンケート調査で「自治体評価の課題」とされる事柄は、その自治体数が多いほど期待どおりに評価が機能していない証左と判断されよう。情報を複眼的に吟味すれば、実態に少しでも近しい分析結果を引き出すことは十分可能である。

近年の自治体評価の実際はどのように捉えればよいのか。ここでは、総務省が全自治体に対して実施した定期調査[13]の結果（平成20・21・22及び25年度の調査結果）を基に、組織管理と評価結果の活用に関する問題意識を経年変化から探り、その実際に迫ってみたい。勿論各自治体の回答は公式的なものであり、真実の披瀝よりもいわゆる厚化粧をした回答となっていることが基本的には考えられる。よって内的妥当性への脅威（Threats of Internal Validity）を前提とし、あくまで大まかな動きを把握するだけに留めたうえで当該情報の分析に解釈を付したい。

（1）課題

図2は、上記調査が取り上げた組織管理業務の中で、評価の活用に「課題」があると認識する自治体の割合をグラフにしたものである。職員の意識改革以外は全体的に右肩上がりであり、評価は組織管理の期待に未だ応え切れていないことが示唆される。特に「評価指標の設定」や

図2　自治体評価の課題



（出所）総務省（2014、2011、2010、2009）に基づき筆者作成

「予算編成等での活用」が高い水準を維持し続けているのは、評価の本質的な問題が手付かず状態であることを浮き彫りにしている。

また「行政評価事務の効率化」の伸びが特徴的で、データは平成20年度と25年度の二つしかないものの、評価事務への負担感は5年間で約1.5倍に増加している。評価システムが期待に応える成果を産出できない結果、それに対する批判が評価事務への負担感へと転化していったとも解釈できよう。

(2) 成果

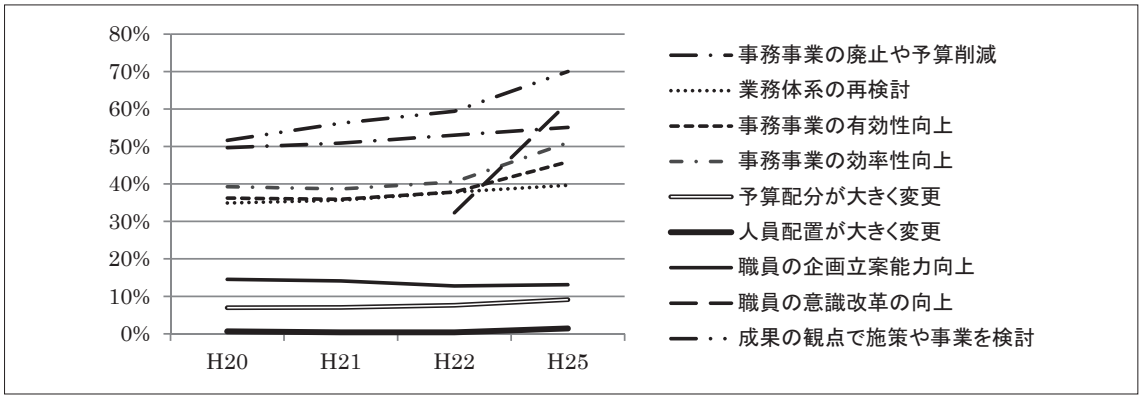
図3は、(1) と同様に当該業務の中で「成果」があると認識する自治体の割合を示したものである。事業の改善・廃止や予算削減にかかる業務については「成果」を認識する割合は高く、右肩上がりで推移している。また「職員の意識改革の向上」については、これもデータは平成20年度と25年度の二つしかないが、5年間で約2倍に増加している。実際に向上したか否かは読み取れないが、振り返りの機会となること自体に評価の意義が見出されていることはわかる。

しかし予算や定員といった経営資源の配分上の成果については低い水準でほぼ横ばいの状態が続き成果の変化はみられず、評価システム全体の貢献度は決して高くないと考えられる。

(3) 予算業務

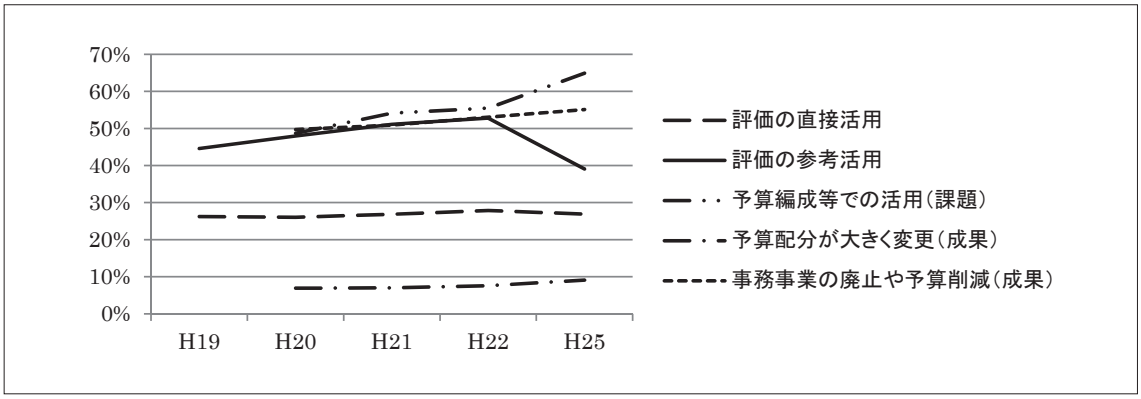
以降では具体の業務について個別に検討を加える。まずは予算業務である。図4は当該業務に

図3　自治体評価の成果



（出所）総務省（2014、2011、2010、2009）に基づき筆者作成

図4　予算業務にかかる評価の活用・成果・課題



（出所）総務省（2014、2011、2010、2009、2008）に基づき筆者作成

かかる「活用の度合い」や「課題」、「成果」に対する認識の割合を示したものである。「評価の直接活用」は低い水準で横ばい状態が続く中、25年度に「評価の参考活用」が急激に減少している。これは近年、予算業務での評価の活用を組織として断念する傾向が出てきたことを示唆するものと解釈できる。

一方で予算削減業務での成果が右肩上がりで推移していることには若干の矛盾を覚える。しかしこれを「予算削減に正当性を付与する名目的な評価結果の活用」の証左として捉えれば決して不自然なことではなく、この言い訳的な活用が評価事務の効率化を求める声を増幅させているとも考えられよう。

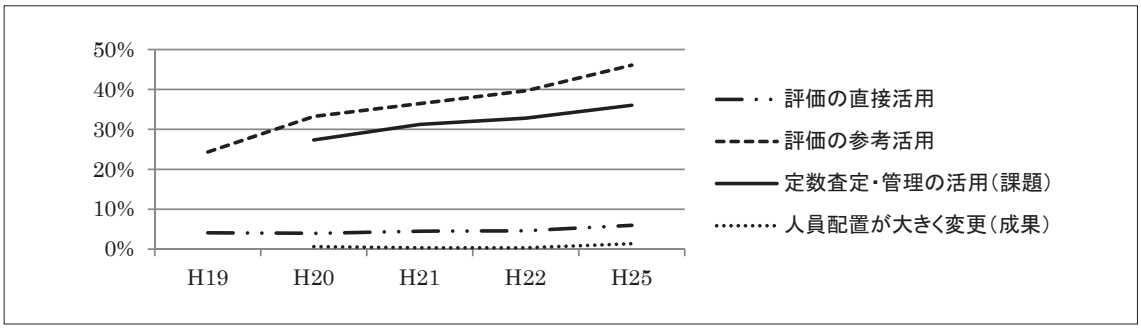
（4）定員業務

続いては定員の管理業務である。図5は当該業務にかかる「活用の度合い」や「課題」、「成果」に対する認識の割合を示したものである。「成果」も低水準で横ばいに推移し、「評価の直接活用」も同様に低い水準を示している。一方、「評価の参考活用」と「課題」はいずれも毎年伸びてはいるものの、図2において後者は最も低い水準に位置している。これらを踏まえれば、定員業務での活用の期待値はそもそも低く、その程度において評価結果は参考活用されていると考えられる。現行システムの評価書の中身に鑑みると、定員管理や査定というよりも事業概要の全体把握という意味での活用と判断するのが妥当であろう。

（5）事務・事業の見直し等の業務

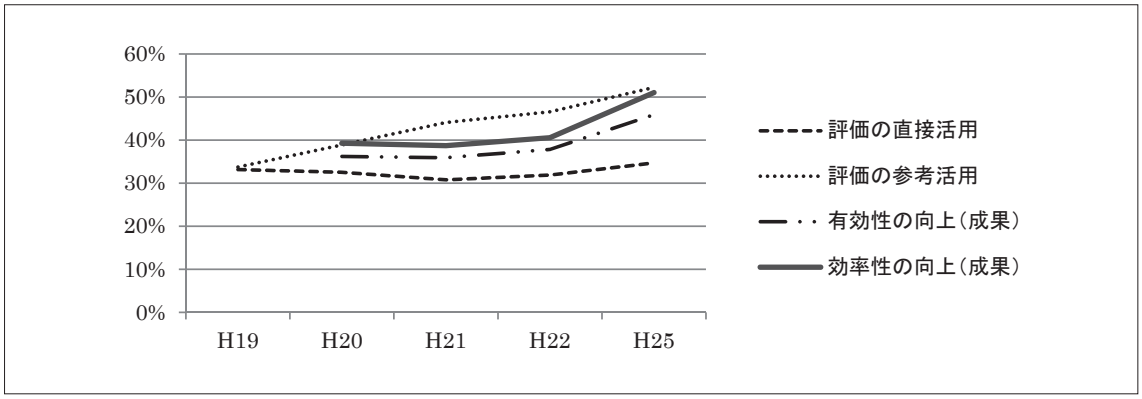
次に事務・事業の見直し等の業務である。図6はその「活用の度合い」と「成果」の認識の割合を示したのである。当該調査では、この業務

図5　定員業務にかかる評価の活用・成果・課題



（出所）総務省（2014、2011、2010、2009、2008）に基づき筆者作成

図6　事務・事業の見直し等の業務かかる評価の活用・成果



（出所）総務省（2014、2011、2010、2009、2008）に基づき筆者作成

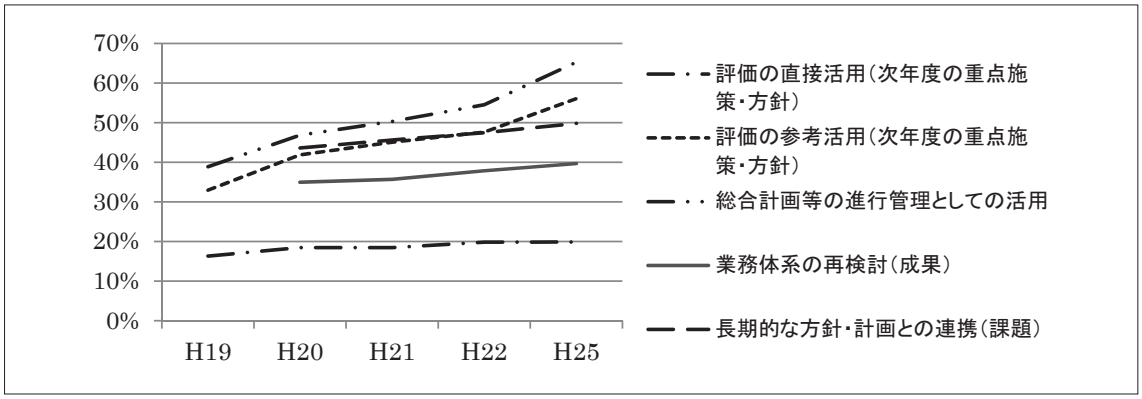
自体が課題として設問の中に位置付けられておらず、評価結果が予算要求前の事業の仕立て直し作業に役立つという前提が暗黙的に成立している。確かに図3においても評価は事務事業の効率性・有効性の向上に貢献しているといえる。しかしながら、それでも「評価の直接活用」は決して高い水準にあるとはいい難く、ほぼ横ばいでの推移が続いている。事業担当者イコール評価担当者であることを考慮すれば、この状態が現行の評価作業ルールでは事業担当者の期待に応える情報を産出できないことを示唆しているとも解釈できよう。

(6) 計画業務

最後は計画に関する業務である。図7は「次年度の重点施策・方針」への「活用程合い」ほか当該業務にかかる「課題」、「成果」の認識の割合を示したものである。

「総合計画等の進行管理の活用」や次年度重点施策や方針への「評価の参考活用」は右肩上がりに推移しており、計画業務における活用は比較的に高い。自治体評価の基本技法が業績測定であることから、施策評価においてその目標の進捗度を確認する手段として有効活用されていることがわかる。一方「評価の直接活用」は低い水準で横ばいであり、また「長期的な方針・計画との連携」を課題とする割合は右肩上がりで増えていることからは、実績値の確認以外に活用し切れていないことが示唆される。

図7　計画業務にかかる評価の活用・成果・課題



（出所）総務省（2014、2011、2010、2009、2008）に基づき筆者作成

7　自治体評価の行方

自治体現場の全国的な傾向からみて、少なくとも自治体評価が劇的に変革を遂げているとはいえず、やはり以前からの課題が未だ解決されていないことが想定される。特に予算業務に関しては第1節で事業部門が主張した「新たな政策統制の仕組み」は具現化されず、予算部門による予算削減の名目的な手法としてのみ存在し、財政部門による伝統的な政策統制が結局は続いているように思われる。

近年、自治体評価を廃止する自治体も実際には散見されるが[14]、評価を廃止する理由を「公表」し、地域の理解を得ることは決して容易なことではない。そもそも評価とは呼吸と同じであり、反射的行為である。自治体の無謬性を受容し、この行為を必要ないとするロジックは今やあり得ず、自らの活動を反芻する何らかの代替案がない限り、地域住民は評価の廃止を容認しないであろう。

では自治体評価はこの活かさず殺さずのまま存続し、事実上の終焉を迎えることになるのだろうか。

残念ながら今の筆者の力量では説得力ある答えを明示することができない。但し敢えて論じることが許されるのであれば、次のような展開が、勿論それらがさらなる撤退か新たな進化への胎動かは判断がつきかねるが、今後においては起こりうるかもしれない。

一つめは評価・分析業務の効率化を伴うシステムの存続であり、事務事業評価の廃止や評価対象の縮小さらには記載事項の簡便化へと向かう動きである。組織管理での活用では評価・分析よりも施策目標のモニタリングに主眼がおかれるであろう。職員にとっては大いに助かる、自発的だが消極的な評価の動きといえる。近年、一部の自治体では米国ボルチモア市の「CitiStat」を例とする業績レビュー会議が導入されている。これによって一部の取組は外部を交えた四半期毎のモニタリングが担保される一方、取組全体をみる現行の枠組はこれまで以上に縮小されていくであろう。

二つめは自発的かつ積極的な評価の動きである。かつてのように複雑な社会的情勢が評価という「政策の窓」を一気に開放し、組織内部から評価システムの変革が台頭するシナリオである。しばしば政治と行政の非公式な対話では民主性・効率性といった基準が軽視され、結果として一部のステークホルダーのみレントを享受する事業が予算化される。仮にこれからの行政がこのような旧態然とした状況に反発するようになれば、今後、客観性と科学性に依拠した議論を展開する仕組みとして行政自らが評価の積極的活用を模索するかもしれない。例えば、いわゆる「口利き」記録制度等と評価が連動するようになることで議員による非公式な依頼を政策提言として公式に分析・評価することが可能となるし、地域住民からも積極的な支持を得ることができよう。

一方、全くの外発的要因による動きも考えられる。中央政府による評価の義務付け強化、それに伴うシステムとしての自治体評価の縮小もしくは連動である。近年、例えば内閣府主導の「都道府県まち・ひと・しごと創生総合戦略及び市町村まち・ひと・しごと創生総合戦略」のように、法令等に基づき自治体が策定する計画に対し、国が数値目標・重要業績評価指標（KPI）の設定を義務付けるケースが増えてきた。評価を義務付ける計画の策定が増えれば増えるほど、それら計画のKPIが自治体評価の目的を事実上代替することになる。その結果、敢えて評価システムを有する必然性は一層低くなっていく。この外発的かつ消極的な評価の動きが三つめにあたる。

しかし外発的であっても自治体自らが積極的に動くシナリオも考えられないわけではない。最後は自治体のオープンデータ化に伴う、評価における競争環境の誕生である。これまで政策情報は行政にほぼ独占され、地域住民は公表された評価結果を信頼するほかなかった。また評価結果を覆す政策情報など容易に入手できなかった。しかし近年は、ICT等の本格的活用を目指し、行政のデータ公開が積極的に推進されている。自治体側も公開によって官民協働による公共サービスの推進や新たな民間サービスの創出を期待する。すなわちこれは自治体データへのアクセス・フリーが評価のイノベーションをも

たらすという発想であり、具体的には民間・個人による自治体評価への自発的参入を意味する。仮に多くの私人による客観的かつ直接的な評価が可能となり、それらの評価結果が公表、拡散されるようになれば、自治体の評価結果は私人による評価結果と比較されるようになろう。その結果自らの評価結果を正当化したい自治体は質の高い評価結果の産出を追求し、積極的に現行評価の改善に取り組んでゆくかもしれない。

終わりに

結局のところ、筆者の主観をかなり振りかざし自治体評価の20年を総括してみた。実効性には懐疑的で事務量の負担感もあるものの、外形的にはPDCA機能の存在や説明責任の徹底を標榜できるし、評価の結果は常に内部で操作できる。それが現在の自治体評価である。要は積極的に評価を改革するメリットはなく、何らかの「政策の風」が吹かない限り現行の自治体評価は、活用されもしないが廃止されることもなく組織内で細々と存続し続けるのであろう。

最終節の4つのシナリオは、組織内外における自治体評価を動かす状況（自発性・外発性）とその動きに対する自治体側の姿勢（積極性・消極性）という2つの軸を用いて予想を試みた。これらは過去の自治体評価の台頭に大きく関係した要因だからである。

以上、学術的貢献をなさない拙稿ではあるが、それゆえに従来までの論考とは些か異なる自治体評価のありさまを記述できたように思う。但し、このような非生産的な論調で終えてしまうのは甚だ不本意であるのも事実である。是非とも実務の世界では、自らの活動のあくなき発展を遂げるため、「肉を切らせて骨を断つ」覚悟で評価を実践してもらいたいと強く願っている。また筆者としても、評価技術の向上を今後も追求しなければならないし、新たな「風」とは何か、その理論的に後押し等々、しつこくも評価研究の世界に居座り、その答えを明示する役割を果たし続けていきたい。

謝辞

本稿の草稿にあたっては査読者の方から丁寧なコメントを頂戴した。この場を借りて御礼申し上げます。

注記

1 第三の目的としては「知識創出」のための評価がある。但し、これは具体の評価結果の活用を意図して実施されるものではない。詳しくはPatton（1997）、Weiss（1998）等を参照されたい。
2 勿論評価の過剰な組織管理志向に対する批判も導入直後から少なくない（山谷（1997、1998）など）。
3 自治体評価とNPMとの連動については例えば山本（1998）を参照されたい。
4 例えば古川・北大路（2001）、上山・井関（2001）、山谷（2002）など。
5 この中では、行政改革推進上の主要事項の第一に事務・事業の見直しが掲げられており、職員参加による明確な目標設定と効果的な進行管理の徹底等、民間手法の導入による行政運営プロセスの改善が求められている。
6 日高（2000）。
7 2000年前後は「行政経営フォーラム」や「自治体有志の会」など評価や行政経営の勉強会に多くの自治体職員が自主的に参画していたことからも、当時の職員の危機意識がうかがわれる。
8 当時の三重県や岩手県での部局への権限移譲や福井県における企画部門での政策予算査定の導入などは、この模索した結果の新たな取組だったと考えられる。なお、この予算査定の新たな動きについては岡本（2003）、竹内（2004）を参照されたい。
9 自治体の持つ「相互参照性」という特性（伊藤（2003））も自治体評価の普及を後押ししたと考えられる。
10 梅田（2001）によれば、当初の三重県の事務事業評価システムは、当初、財政や組織面での改革というよりも「行政風土の改革」の推進が主たる目的であったとされる。
11 例えばWeiss（1998：4）は、評価の鍵的要素の第一に「体系的な調査分析（systematic assessment）」を挙げている。
12 総務省（2014）によれば、評価を導入している団体

のうち、約4割が内部評価に加えて外部評価も実施している。

13 本稿では、総務省の定期調査で使われる行政評価は自治体評価と同意語として取り扱う。

14 総務省（2014）によれば、平成24年度までに廃止した自治体数は計52団体にのぼる。

参考文献

伊藤修一郎（2003）「自治体政策過程における相互参照経路を探る：景観条例のクラスター分析」『公共政策研究』第3号、79-90頁

今井照（2001）「基礎的自治体における「評価」」『都市問題』第92巻第7号、53-65頁

上山信一（2001）「行政評価の本質と科学革命：我が国自治体の行政評価を手掛かりに」『公共政策研究』第2号、70-81頁

上山信一・井関友伸（2001）『自治体再生戦略：行政評価と経営改革』日本評論社

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的過程と展望」『日本評価研究』第1巻第2号、69-77頁

大住荘四郎（2002）「基礎自治体におけるベンチーマーク導入の意義」『NIRA政策研究』第15巻第6号、38-44頁

――――（2005）「戦略的な自治体経営と行政評価の意義」『NIRA政策研究』第18巻第17号、56-64頁

岡本全勝（2003）「予算査定の変容」『地方財務』12月号、87-103頁

小野達也（2001）「地方自治体の行政評価システム課題と成功への条件」『富士通総研研究レポート』、第117号

――――（2008）「都道府県の評価の10年：定量評価のこれまでとこれから」『日本評価研究』第8巻第2号、19-38頁

北大路信郷（2001）「自治体における政策評価導入の意味」『都市問題研究』第53巻第10号、77-88頁

窪田好男（2008）「公共政策学・政策評価論・日本型政策評価」『日本評価研究』第8巻第1号、59-71頁

佐藤徹（2001）「自治体総合計画と連動した施策評価システムの構築に関する基礎的考察」『日本評価研究』第1巻第2号、53-60頁

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』、第140号

高寄昇三（2000）「行政評価システム導入の課題」『会計検査研究』第21号、49-61頁

竹内直人（2004）「福井県庁における予算編成過程の変化」『地域公共政策研究』第10号、10-20頁

田中啓（2003）「諸外国における評価の現状と日本における評価能力の必要性について」『NIRA政策研究』第16巻第5号、4-13頁

――――（2005）「米国における自治体経営とベンチマーキング」『NIRA政策研究』第18巻7号、31-43頁

田渕雪子（2005）「地方自治体における行政評価の現状と今後の展望」『NIRA政策研究』第18巻第7号、65-73頁

西出順郎（2003）「地方自治体における行政評価の実際と新しい評価者モデル：エンパワメント型行政評価者」『日本評価研究』第2巻第3号、179-187頁

――――（2005）「行政評価の再構築：理論着眼型評価志向の確立に向けて」『日本評価研究』第5巻第1号、15-25頁

――――（2012）「評価者は存在するのか：地方自治体の評価システムを中心に」『日本評価学会春季第9回全国大会発表要旨集録』、97-100頁

日高昭夫（2000）『自治体職員と考える政策研究：分権時代の新しい政治行政作法』ぎょうせい

古川俊一・北大路信郷（2001）『公共部門評価の理論と実際』日本加除出版

真山達志（2001）「自治体における事業評価導入の多面的意義」『会計検査研究』第24号、45-53頁

南学編著（2003）「行政経営革命：「自治体ABC」によるコスト把握」ぎょうせい

森田祐司・中嶋崇（2004）「行政評価Next Step：真のマネージメントツールを目指して」『日本評価研究』第4巻第1号、87-96頁

山本清（1998）「地方自治体の行政改革戦略と政策評価（上）（下）」『地方財務』7月号116-134頁、8月号59-70頁

山谷清志（1997）「自治体改革における「政策評価」の必要性：地方分権、行財政改革を実りあるものにするために」『地方財務』12月号、1-10頁

山谷清志（1998）「我が国の政策評価の現状分析」『行政管理研究』第84号、52-64頁

山谷清志（2002）「政策評価とNPO：もう一つの実験」『公共政策研究』第2号、82-95頁

――――（2007）「政策評価とその改善：実践と研究の交錯」『評価クォータリー』創刊号、2-13頁
Patton, M. Q.（1997）. *Utilization-Focused Evaluation*, third edition, Thousand Oaks: Sage Publication.
Weiss、H. C.（1998）. *Evaluation*, second edition, Upper Saddle River: Prentice Hall
総務省（2014）「地方公共団体における行政評価の取組状況（平成26年3月25日公表）」総務省ホームページ
http://www.soumu.go.jp/main_content/000278815.pdf（2015年12月 1 日閲覧）
――――（2011）「地方公共団体における行政評価の取組状況（平成23年3月16日公表）」総務省ホームページ
http://www.soumu.go.jp/main_content/000156645.pdf（2015年12月1日閲覧）
――――（2010）「地方公共団体における行政評価の取組状況（平成22年3月16日公表）」総務省ホームページ
http://www.soumu.go.jp/main_content/000156645.pdf（2015年12月1日閲覧）
――――（2009）「地方公共団体における行政評価の取組状況（平成21年3月31日公表）」国会図書館ホームページ
http://warp.da.ndl.go.jp/info:ndljp/pid/3487671/www.soumu.go.jp/click/jyokyo_20081001.html（2015年12月1日閲覧）
――――（2008）「地方公共団体における行政評価の取組状況（平成20年3月5日公表）」国会図書館ホームページ
http://warp.da.ndl.go.jp/info:ndljp/pid/3487671/www.soumu.go.jp/click/jyokyo_20071001.html（2015年12月1日閲覧）

（2016.6.24 受理）

Looking back on Performance Evaluation System in Local Government: Two Decades of the Non-utilization nor Discontinuance

Junro Nishide

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

Abstract

The purpose of this article is to look back on the history of 20 years of performance evaluation system in Japanese local government from a viewpoint of the research and practice, and try to forecast the future of the evaluation system. First, the reasons why the utilization on the evaluation system mainly focused on internal resource-management were explained from environment surrounding the local government around the year 2000 when local government evaluation gained power. Second, the problems of the evaluation system that such factors caused, the prescription for solving them and their limits were discussed. Final part indicates discussions on the recent trend of the evaluation system and the implications for the future vision of the evaluation system.

Keywords

Local Government, Performance Evaluation, Organization Management

【総説】

地方分権改革と財政危機の自治体評価
ー20年のレビューからー

山谷　清志
同志社大学 政策学部
kyamaya@mail.doshisha.ac.jp

要　約

地方自治体が評価に取り組んで20年、その呼び名は地方分権改革時代の政策評価、自治体財政破綻・財政危機を意識した行政評価・行政経営改善と、そのときどきの政治課題に対応していたが、実務の現場では公共事業評価、施設評価、指定管理者評価、地方独立行政法人評価、学校評価、病院評価などの専門性が高い評価が次々に加わってきた。さらに第二次安倍内閣の2014年、地方創生がアジェンダとしてクローズアップされたが、そこには重要な思考ツールとしてPDCA（plan-do-check-act）サイクル、KPI（Key Performance Indicators、重要業績指標）が組み込まれて、これを「評価」と言う人も多い。本稿は20年にわたる地方自治体評価の実践をレビューしながら、何を目的に、どのように行ってきたのか整理を試みる。

キーワード

政策評価、行政評価、目標管理型業績測定、PDCA、KPI、「評価のガラパゴス化」

はじめに

日常用語で普通に使われる評価ではなく、学術的専門用語として意識される「評価（evaluation）」が、日本の内政における実務に登場するのは、1990年代半ばの二つの改革、中央省庁改革「橋本行革」と地方分権改革がきっかけである（なお政府開発援助では海外経済協力基金が1975年から、また外務省が1981年から評価を開始していることに留意すべきである）。

日本の地方自治体は1990年代半ばから、さまざまな領域で、多くの評価を順次採用してきた。それらを時系列で並べると、一番目に政策評価、二番目に行政評価、三番目に多くの分野における多種多様な評価、そして四番目に「地方創生」に付随する評価の試みである。そして、これら4種の評価は目的が違い、方法が異なり、評価結果の公表方法が別々であったため、一般市民は簡単に理解できなかった。

ただ、同じように1990年代半ばから日本でも多用されるようになったアカウンタビリティ概念を使って、これらの評価を整理すると分かりやすいかも知れない。たとえば、地方自治体の評価ブームの先駆けになった政策評価は、政策がその目標をどれだけ達成したのかを調べる方法である。その先にあるのは、こうした政策を作った首長の政治的正統性（legitimacy）である。この政策評価が注目されたきっかけは、1993年の衆議院と参議院の地方分権推進決議、1995年の地方分権推進法であった。長年の懸案であっ

日本評価学会『日本評価研究』第16巻第1号、2016年、pp.31-45

た中央（府省）と地方自治体との関係を抜本的に見直す地方分権改革が、「地方自治体も『政策官庁』として自立すべき」と考える運動と連動していたからである。地方分権の結果、地方自治体は国の府省が作った政策の実施機関（従属的下請）でなく、自ら政策を作る主体的「政策官庁」になる。そうであれば、政策の見直しや効果の評価を行う仕組み、つまり政策評価が必要だと多くの地方自治体が認識したのは当然である。この認識に理論的裏付けを行う場として日本公共政策学会が1996年に設立され、研究者だけでなく、多くの地方政治家（首長や地方議員）、そして地方公務員が参加したのも、政策評価に地方自治体が注目した背景にある。

行政評価はこの政策評価に比べると、きわめて実務的で地味である。行政評価が注目されるようになったのは地方自治体の財政危機、とくに北海道夕張市の財政破綻（夕張ショック）であって、ここには「実務上の事務事業編成や予算などの行政のプロが、日常的に行っている内部管理の地味な作業に深くコミットすることが、財政破綻を回避できる」と思う信念があった。しかし、その一方で役所の予算書や事務事業評価書に興味を持つ一般市民は少ないので、行政評価は地味なテーマになったのである。ただし、市民にとって痛みを伴うこともある。それはこの行政評価がコスト削減、予算カットにつながることが多いからである。それでも行政評価に市民が反対しなかったのは、行政評価が地方自治体の経営破綻回避策であり、ムダ・浪費の削減が健全な行政を実現する方法であると合意されていたからである。

その後、評価が氾濫する時代が来る。介護保険、地方独立行政法人制度（公立病院・公立大学・研究試験機関）、学校、公共施設などに対して、新公共経営（NPM:New Public Management）の考え方を反映した指定管理者制度手法やPFI（Private Finance Initiative）、PPP（Public Private Partnership）などの手法が導入され、ここに多くの評価が付随していた。まさに「評価」が氾濫したのであった。理由は簡単である。仕事を繰り返す中で一度見直すべきだという常識がまずあり、行政サービスは市民からどう思われているのか一度考えてみるべきだというこれまた常識的な反省もあった。もちろん教育や医療、公共事業などの専門家・専門職の人たちは、学会や研究者同士での議論を重ねてきているので、そもそも自己評価に積極的な気分を持つ。

最後に登場したのは、第二次安倍政権の「地方創生」（2014年）に付随する業績評価関連のPDCAとKPIである。これらはアカデミックな議論というより、現場主義的な品質改善にそのルーツを持っている。有名な「トヨタのカイゼン」に近い。そして目標達成度を見せられるトップは現場の実際を観察する必要は必ずしもなく、重要指標を見れば必要な変化が分かるようになっている。

以上の四つの大きな流れの中で、目的、価値観、判断規準、背景の学問や専門が異なる「評価」が20年の間に次々と登場してきたが、現場の公務員、政治家、そして市民が理解できているとは言えない状況がある。本稿では、これらの評価群を整理し、その考え方の本質を見極め、市民自治にとってどのように貢献できるか、考えてみたい。

1　地方分権改革と「カイゼン」

地方自治は戦後日本の民主化改革として成果が出ているかどうかいささか疑問は残るが、長い間改革のアジェンダに乗り続けた。ただし、地方自治においては、1947年に制定された地方自治法に見られるように、常にマクロの民主主義改革と、経営の能率を追及するミクロの改善が混在してきた（なお、政策評価が注目された1997年頃、経済学の影響で「効率」と呼ばれることが増えたが、実務の認識では能率も効率も同じである）。

第一条　この法律は、地方自治の本旨にもとづいて、地方公共団体の区分並びに地方公共団体の組織及び運営に関する事項の大綱を定め、併せて国と地方公共団体との間の基本的関係を確立することにより、地方公共団体における民主的にして能率的な行政の確保を図るとともに、地方公共団体の健

全な発達を保障することを目的とする。（傍点は筆者）

その結果、地方自治法の趣旨に従う改革と言いながら、ある人はさらなる民主化に向けた改革をイメージし、別の人は技術的な経営効率追求手段を考える。地方分権改革にもそれが見られ、日本国憲法に代表される戦後「民主化」改革の流れに地方分権改革を位置づけて考える人と、技術的な能率・効率を追及する改革ととらえる人が混在していた。それが、評価の考え方にも影響したのである。つまり、さらなる民主化を地方分権に重ね合わせて見る人たちは政策評価の信奉者になり、経営効率を唱える人たちは「カイゼン」のエキスパートをめざすのである。

ところで、行政評価の背景にあり、推進力になったのは20世紀末の国際的な新自由主義の潮流、新公共経営（NPM）運動で、その運動が政治的なアジェンダになったことが大きい。つまり、NPM改革を信奉する首長が、議員や公務員の給与カット、民営化、議員・職員数の削減などを進める「改革派」と賞賛され、選挙戦を有利に戦う場面が増えたのである。そうした削減やカットのテクニック、ツールを提供するシンクタンクやコンサルタントも存在した。

たとえば行政経営品質向上運動である。この運動は当初、伝統的な生産現場での品質改善運動に近かった。有名になったのは東京都三鷹市で、ここでは社会生産性本部の支援を受けて行政経営品質評価基準を開発し、その後全国的注目を浴びた（1999年6月ごろ）。ほぼ同時期、岩手県、三重県もこの品質改善運動に取り組んでいたし、また当時日本一人口が多い「村」として注目を浴びた岩手県滝沢村（2014年に市制移行）は、2002年に政策評価とともに行政品質向上を行政改革の二つの柱にした。この品質向上「運動」の活動内容は、QC（Quality Control）サークル的な意識改革活動である。対象になった自治体の経営の質はこの活動によってベンチマーキング・評点を使って比較され、それがまた自己診断されるが、最終的には地域住民が評価するので、住民満足度調査の要素が強かった。ここでは「住民＝顧客」と置き換え、民間経営のアナロジーで自治体経営を考えたい人が、経営コンサルタントに委託してカイゼンの処方箋を求め、試行錯誤した。21世紀になってから評価と同義でPDCAが多用されるようになった背景には、こうした背景がある。

ほぼ同時期に流行した「国際標準化機構」のISO（International Organization for Standardization）の規格取得も、地方自治体がその独自性を発揮する戦略に使われた。ISO9001の品質管理と品質保証、環境保全のISO14001は、外部の権威ある民間機関、国際標準化機構の認証を受けることを前提に、その準備作業に経営見直し、組織改革、「カイゼン」がセットで含まれており、また事後的にも「カイゼン」の継続をチェックされるため、経営の品質について認証を定期的に受けることになる。中央省庁からの指示を待って、言われたことをこなすスタイルを辞めて、自らのイニシアチブによって改革・改善を進める姿が地方分権と重なったため、ISOを取得する地方自治体が増えていった。なお、このISO取得が政策評価から経営改善型の行政評価に地方自治体が向かう隠れた背景であったことは知られていないが、PDCAを評価と重ねて見る発想にその痕跡が残っている。

同じような「カイゼン」運動に、『日経地域情報』誌が主催した「六七〇市･東京二三区の行政革新度」(1998年9月7日no.302）がある。行政改革に関して熱心な地方自治体を紹介し、ランクづけて上位者を表彰することが中心であるが、このランクの上下には全国各地の地方自治体が重要な関心を持った。ランクづけの項目は透明度（企業会計方式・オンブズマン制度・外部監査と民間人登用）、参加度（構想策定に一般市民参加・NPO支援の制度）、利便度（ワンストップサービス･自動交付機設置･図書貸出し広域化と閉館時間延長）であった。この当時の総合評点でAAA（スリーエー）を付けられたのは三鷹、松本、川崎の3市であり、多くの自治体は3市の改革を競ってコピーした。

このように、地方分権改革の中で経営改善のフィーバーが起き、その文脈で地方自治体の行政評価が注目されたのである。したがって、政策評価が本来の政策の見直し、政策の優先順位の変更、政策の終了・廃止の検討手段ではなく、

経費削減や健全な経営マネジメントのツールであると誤解されるのも無理はなかった。政策評価も行政評価も同じ文脈で、同じ目的で考えられ、良く言えば首長の経営戦略のPR、しかし実態は組織いじりやコスト・カットの手段になり、「乾いた雑巾をなおも絞る」経営が美徳と唱えるような、情緒的なキャッチフレーズを各地のシンポジウムで耳にすることも少なくなかった。

ふり返って考えると、この経営改善ブームが起きたのがバブルの名残がまだ残っていた時期だったことが問題であった。財政的にも精神的にも余裕がある気分の中で経営改善が主張されたのである。したがって、無駄削減の「カイゼン」運動に数千万円の予算が組まれることも珍しくなく、経営コンサルタントにとっては新たなビジネスチャンスになった。また、「カイゼン」の業務委託に多額の予算を用意できた地方自治体は「カイゼン」に熱心だと自治体経営の専門誌で賞賛され、他方、予算を用意できない小さな地方自治体は「居眠り自治体」とレッテルを貼られ、その後の市町村合併の対象になっていった皮肉な現実もある。

もちろん組織内部の改善の話だったため、一般市民にとっては直接の利害も関心もなく、地域の新聞で報道されれば、「何か新しいマネジメント改革をしているのか」、その程度の認識であった。まれに、勉強熱心な一般市民が行政評価外部有識者会議に公募や自己推薦で参加しても、何が何だか分からない膨大な資料を見せられ、意見を求められることもしばしばあった。

地方自治体がその組織経営の健全さ、グッド・ガバナンスを競い公表することは重要なことだが、組織の内部管理に無関係な外部市民にとってはあまり関心がない。ガバナンスが悪く自治体倒産の憂き目に遭う状態でもない限り、健全な行政経営は当たり前のことなので、とりたてて誇ったり自慢したりすることではない。民間企業が投資家に公表する経営健全の話とは次元が違うのである。

ただ、2001年4月に就任した小泉純一郎首相が新設した経済財政諮問会議の「今後の経済財政運営および経済社会の構造改革に関する基本方針」（骨太方針2001）が、地方財政計画の歳出の見直しを求め、地方財政の健全化をすすめ、地方交付税を見直し（実際は減額）、その後「三位一体の改革」が市民の意識を自治体改革や業務「カイゼン」に誘導したため、一般市民は「削減・カット＝カイゼン」の雰囲気に馴らされた面はある。この雰囲気の中で、政府機能を政策の立案（舵取り）と実施（漕ぐ）に分けて考えるOsborneとGaeblerの*Reinventing Government*（1992）も『行政革命』として翻訳され（1995年）大きな影響力をもつが、改革の対象は「舵取り」でなく「漕ぐ」（組織活動）の省エネ方向に収斂したのである。

その意味で20世紀末から21世紀当初数年間のブームは、ミクロの「カイゼン」が大きなマクロ改革であると勘違いさせられた、残念な時期だったのかも知れない。改革が民間コンサル「お任せ」で、実態としてそれが巧く行っているかどうかは不明で、しかし地方自治体本来の民主的な視点から言えば何か良くなったかと言えば、実は何も変わっていない、そんな時代であった。

それではなぜ、こうした勘違いになったのであろう。それは先の地方自治法「民主主義と能率」に戻って考えてみる必要がある。つまり、経営スキルの技術的な話が、民主主義とは関係ないはずであるのに、大きなPRの後にブームになったので、市民から見れば反対意見をはさみにくくなったのである。もっとも、こうした改革の当事者は良かれと思って善意で行っていたのであり、それも勘違いを招く原因になった。

2　政策評価の登場

ところで、地方自治体の評価ブームは、何もないところで突然始まったものではない。長年の研究や試行錯誤の延長にあった。たとえば旧・自治省事務次官通知「地方公共団体における行政改革推進のための指針について」（1994年10月7日）は、地方自治体に事務事業の見直しを求めており、多くがその方向で取り組んでいたことがわかる。

「第2　地方公共団体における行政改革推進のための重点事項について」、「1　事務事業の

見直し　地方公共団体が、限られた財源の中で新たな行政課題や社会経済情勢の変化に的確に対応していくためには、事務事業について絶えず見直しを行い、喫緊度の高いものを選別し、効率的な事業の実施を図る（中略）」

もっとも、その実際の運用としては行政改革委員会を設置して、組織機構・定員の見直しに重点化した従来型の行革に終わった地方自治体も少なくない。政策、事業の評価がいかなるものであるのか、事前に十分研究していなかったからであろう。

ただし、事前に研究を重ねていた地方自治体も存在した。兵庫県川西市が1992年度から地方自治資料センターと共同研究で試行・導入していた「行政評価」が知られており、また神戸市の神戸都市問題研究所「自治体における行政評価システム」の提言もある（1993年3月）。

その後の展開の中で、日本の評価システム導入のトップランナーとして記憶された三重県庁は、1997年に「事務事業目的評価表」作成に着手し、また「平成10年度行政システム改革案」（1998年）では、改革の柱のひとつに事務事業の見直しを置いていた。窪田（1998年）によると、三重県の「事務事業目的評価表」のポイントは、次のとおりである。

①公共関与の妥当性
②県関与の妥当性
③対象の妥当性
④意図の妥当性
⑤手段の妥当性
⑥出先機関の意見の検討および本庁と出先機関の役割変更余地
⑦他部門・他事務事業との機能重複及び関連性
⑧成果向上余地・成果向上のための問題点
⑨コスト（予算・所要時間）の節約余地やコスト節約のための問題点
⑩事務事業のポジショニング（予算を更に投入してもこれ以上成果は出ない）

三重県庁がはじめに構想していた評価（政策評価）がマクロの政策や戦略を前提にしていたのに対し、実際の事務事業目的評価表の基本のコンセプトは県庁内での行政運営の見直し、現場の活動のチェックの姿で現れた。

三重県庁以外にも政策評価導入に向かっていた地方自治体があった。北海道庁は「時のアセスメント」を1998年から99年にかけて「政策アセスメント」に展開し、岩手県庁と秋田県庁も1998年頃に政策評価を採用した。また、静岡県庁は静岡総合研究機構と共同で1997年頃から「業務棚卸」と名付けた政策を棚卸しする仕組みを実践しはじめていた。静岡県の改革項目は、課の廃止で組織をフラット化する、組織のリーダーシップを強調する、目的＝手段の連鎖構造の重視、アウトカム評価の採用などであった。なお、早期に政策評価として条例を制定した地方自治体は北海道（政策評価条例、2002年3月29日）、秋田県（政策等の評価に関する条例、2002年3月29日）、岩手県（政策等の評価に関する条例、2003年10月9日）であった。

当時の社会状況を想い起こすと、社会全体で改革や新しい試みの柱に評価を入れ込もうとする「評価文化」が生まれはじめたように思われる。あるいは、そうした雰囲気が21世紀を目前にした日本社会にあった。「今後の地方教育行政の在り方について」（1998年9月）は以下のように言う。

「各学校においては、教育目標や教育計画等を年度当初に保護者や地域住民に説明するとともに、その達成状況等に関する自己評価を実施し、保護者や地域住民に説明するように努めること。また、自己評価が適切に行われるよう、その方法等について研究を進めること。」

また、介護保険制度の導入時の議論には、要介護認定、サービス主体・事業者の選定、地方自治体の能力などをめぐって、評価に関する話題が多く含まれていた。あるいは「健康日本21」（21世紀における国民健康づくり運動）によって厚生省（厚生労働省）は、科学的評価に基づいた目標の設定と計画実施の分析評価を求めていた。

もちろん、この評価を肯定し積極的に受け入れようとする社会的雰囲気と、地方分権運動が結びつき、それを改革派の首長が取り入れ、首長のリーダーシップによって自治体の改革運動

が進んだ事情は、既に述べたとおりである。その結果、評価に関心を持つ首長は「改革派」と呼ばれるようになったのである。

改革を標榜する首長が採用する「政策評価」は、都道府県や市町村の総合計画・基本構想の中に評価の仕組みを入れるスタイルで着手するところが多かった。政策局、政策部、企画政策課など組織名称に「政策」がある部局が事務担当で、外部の有識者による委員会（策定委員会・推進委員会・市民委員会など）の審議を経て、議会の承認を得て成立する総合計画に政策評価の仕組みを入れることが全国で流行した。とくに20世紀末、あるいは21世紀初めに総合計画が更新期に入った都道府県はこうした方法を採用した。また、地方自治法第2条第4項で総合計画の基本構想の策定が義務づけられていた市町村も、その計画の中に政策評価のメカニズムを組み込もうとする場合が多かった。ここでの主な評価手法は、計画や構想の目的に達成指標を付けて「目標」とし、数年後にその達成度を測定する方法、いわゆる業績測定（performance measurement）である。ただし、こうした測定は数年に一度のイベントになり形骸化する。

3　行政評価としての普及

もう一つの評価として登場したのは、事務事業を対象にした行政評価である。計画や構想を対象にする政策評価と比べ、事業担当課が所管する事務事業（予算の単位）を対象にしているため、実務担当者の迷いは少なかった。全庁的に予算編成作業に活用する作業に組み込まれ、総務部の評価担当グループと財政担当が結合した部分が行う「行政評価」、実態は事務事業評価とそれを束にした施策評価として一般化する（この施策評価は政策評価理論で言う「プログラム評価」ではない）。ただしこの事務事業評価・施策評価は特定の手法、方法論に導かれたものではなく、公務員であれば誰でもできる「書類に書き込む」記入方式である。この簡便さと、財務会計部局を中心とした予算査定に活用したいニーズが、事務事業評価方式を多くの地方自治体に普及させる。

もちろん、事務事業を担当する課の単位で事務事業担当者が自ら書き込むため、内部管理事務になっていく。それを「経営評価」と呼ぶ地方自治体も現れる。内部管理事務なので外部有識者からなる行政評価委員会は必要なく、したがって委員会を設置しない地方自治体も増えた。さらに事業仕分けの資料にこの事務事業評価書が使われたが、仕分けの過激さに懲りた地方自治体では、事務事業評価結果を使って事業を廃止・削減するメカニズムをやめた（廃止ではなく「抜本的見直し」を使う）。したがって、この事務事業評価は公務員にとっては無害な「恒例年中行事」化し、やがて煩瑣な手間を嫌がる地方自治体では、このメカニズムすら次第に省略される。

こうして、政策評価として評価を導入した地方自治体も、まだ導入していなかった地方自治体も、評価の目的を予算と組織活動を見る測定活動に変え、それを行政評価と呼び、やがて年間の予算編成サイクルにあわせた経営評価にしていく。この経営評価になる段階でマネジメント・スキルが求められ、場合によっては2014年度から公会計制度を入れた愛知県庁のように、管理会計の仕組みに近づいていく。

こうした流れとは別の試みを考えた地方自治体もあった。市民参加型の自治体評価をめざす動きである。たとえば、滋賀県草津市は「住民協働型行政評価」を構想し、2004年6月から試行した。最初に、全国の地方自治体に行政評価の外部評価をどのように行っているのか照会したが、照会結果が【表1】である。各地にある外部評価委員会は、その形式上は市民参加であるが、全市民を対象にした評価の会議は物理的に無理である。そこで公募委員という形で市民に手を上げてもらうか、行政評価担当課が選んだ市民委員に参集してもらい、議論をする方向が模索された。ただ、各地の地方自治体評価の実際は、「恒例行事」の会合を年に数度開催して「ご意見を伺う」方式になってしまう傾向があり、参加本来のあり方をめざして苦戦していた状況が浮かび上がった。参加、参画、協働の実効性をどのように上げるかが課題になっていたのである。

表１　草津市が2004年当時調査した外部評価委員会の状況

自治体名	委員会名	委員数	市民数	備　　考
名古屋市	行政評価委員会	10	0	1,826 事業で、委員１人あたり 100 時間
京都市	事務事業評価委員会	6	0	学識経験者のレビュー
所沢市	外部評価会議	3	0	586 事業のうち 33 事業を教授、公認会計士、行政経営アドバイザーがレビュー
杉並区	外部評価委員会	5	0	５政策、19 施策を有識者と区内在住専門家がレビュー
芦屋市	行政評価委員会	8	0	132 事業を有識者がチェック
中野区	外部評価委員会	5	3	－
日向市	行政評価委員会	8	－	専門家と「まちづくり 100 人委員会」がレビュー
横須賀市	まちづくり評価委員会	5	11	専門家と市民委員 11 人（市民委員のうち公募が６人）が評価表をチェック
板橋区	行政評価委員会	6	－	80 施策から 11 施策、135 事務事業から 29 施策を弁護士、公認会計士、行政評価のプロ、公募区民からなる委員会がチェック
滋賀県	行政経営改革・施策評価委員会	20	0	学識経験者が全般についてレビュー
三重県	設置しない			ＮＰＯ「評価みえ」に外部委託

（出所）草津市行政評価市民委員会の調査資料（2004年9月）をもとに著者作成。

市民参加の理想はあるが、現実問題として、どのような市民に、どんな形で参加してもらうのか見えにくいのである（なお2015年現在も「草津市協働のまちづくり・市民参加推進評価委員会」として活動している）。

一般に評価を導入する目的として掲げるのは、欧米でも日本でも、行政のアカウンタビリティの追及、（政策や事業を担当する）組織が行う事業マネジメントの支援、そして政策の内容に関わる専門分野に対する知的貢献の三つである。実はこれに加え、アメリカの地方政府やODAの実践では2000年頃既にエンパワーメント評価、すなわち市民に評価を学習する機会を提供した上で参加型評価をもとめる方法や、市民と行政がコラボレーションする協働型評価の方法が実践されていた。評価を通じた市民教育である。この教育によるエンパワーメントや協働が、評価の四つめの目的として指摘されていた。

しかし、2004年当時の日本では時期尚早で、一部の研究者や専門家を除き、エンパワーメントや協働と評価を結びつける発想は少なかった。議会制民主主義を前提にして、市民と議会が協働評価する場面もなかった。むしろ、参加型評価が首長のリーダーシップ強化策として使われた。あるいは公務員の意識改革ツールとして使っても、さらにもう一歩踏み込み、評価への市民参加による市民のエンパワーメント、市民が政策内容に関わる専門知識を得る、こうした理想を実現するところまで力が及ばなかった。結果として、行政が主体となって市民の代表と称する人たちを招き、評価担当公務員が夜の6時から9時過ぎまでサービス残業でつきあい、あるいは土曜日や日曜日に市民評価会議を開催する、その中で公務員が消耗する場面が多く見られたのである。

それでも市民参加の建前が大事なので、行政の組織マネジメントの手法になった評価、内部評価に市民参加を求める試みが日本の各地で、文字通り試行錯誤された。この試行錯誤には3つの難しさが出現した。一つめは評価理論と政策・事業の内容を区別して議論してもらう難しさ、二つめは政策をマネジメントすることについて理解してもらう難しさで、本当の市民参加型政策評価は組織活動の効率や節約ではなく政策の有効性を問う議論をして欲しいが、そうならなかった場面が多かった。三つめは、いつ集まるのか（平日の夜か週末の日中か）、誰に参加してもらうのか（主婦・現役をリタイアした人･大学

生･公務員OB）などきわめて実際的な難しさであった。もちろんこれら3つは簡単に解決できることではなく、長い間各地の地方自治体が苦しみ続ける課題になっていく。

また、小規模自治体は別の課題を抱えていた。人口が3万人以下の小規模基礎自治体はもともと職員数が少ないため、行政評価に取り組むこと自体が物理的に難しかった。たとえば行政評価の流行が始まった1999年から15年後の2014年、総務省自治行政局行政経営支援室「地方公共団体における行政評価の取り組み状況に関する調査結果」（2014年3月25日）によれば、小規模基礎自治体における行政評価の導入は34.9％にすぎなかった。その理由としてまず考えられるのは、以下のような事情であろう。

小規模自治体は平成の大合併後の削減を経て、職員定数がスリム化している（総務省の「集中改革プラン」2005年でも削減を強く求めている）。合併した場合、いままで通りの職員数という訳にはいかず、当然のように定員削減を求められたからである。しかし、合併後の面積は以前より広くなり、その広い面積の中に基礎的行政サービスの対象市民が点在するため手間がかかる。この職員の削減は難しい。介護や教育、福祉の分野では国が法令で職員配置を決めているからである。むしろ、高齢化や保育の面で手を取られる場面は以前よりも増えているため増員を求める圧力が強い。そこで削減対象になるのは一般行政職、とくに総務系の人員である。そして小規模自治体では総務系人員の削減がボディブローのように効き、しわ寄せが行政評価に来ていたのである。2～3人の担当者で膨大な事務処理を迫られる評価は、手間がかかる無意味な苦行、このイメージが小規模自治体に広がった。

4　評価の氾濫

このように日本の評価ブームは、1990年代末から3回来た。再度確認すると、最初は政策評価、次に行政評価であり、そして第3回目はさまざまな分野で評価が注目をうけて実施され、評価の氾濫を招いたブームである。ただし、3回の評価ブームはそれぞれ重なっており、しかもこの三つを全部行政評価として一つに括る地方自治体もあって、一般市民にはわかりにくい。

わかりにくい理由は、もう一つある。前述した小泉内閣の「骨太方針2001」である。小泉改革は「20年遅れのサッチャリズム」と言われたように、サッチャー改革の思想を実践するNPMの考え方を数多く取り入れた。この動きに影響を受けた地方自治体は、さまざまな実践手法を展開した。たとえばコスト意識（削減）の徹底、県や市の関与の妥当性（どこまで関与するのか）、行政資源の最適配分（無駄はカットして緊急性がない事業は後回し）、市民参画型行政運営（トップダウン型アプローチに基づく市民への情報提供・ワークショップとアンケートを通じた市民参画）、生活者や地域の視点に立った機動的効率的行政システムの確立、市民満足度の向上、県行政システム改革大綱に基づく評価システム（愛知県）などである。すべて一種の精神論的言説で、それが良く言えば評価の「百花繚乱」である。その代表が以下である。

- 民間資金等の活用による公共施設等の整備等の促進に関する法律（1999年）。いわゆるPFIで、民間の資金、経営能力を借りて公共施設を建設、運営する。PFI事業の評価が選定基準になる。なお、これを拡大した考えがPPP（Public Private Partnership）である。
- 地方独立行政法人制度（2003年）。地方独立行政法人法に基づく評価は、基本は首長が出す中期目標を独立行政法人が中期計画として設定、それを中期目標期間中にどれだけ達成したのか、その業績評価が中心。ただ、財務諸表、コスト情報などを、公会計制度を一般行政活動に導入した地方自治体の行政評価と同じように使用する。
- 競争の導入による公共サービスの改革に関する法律（2006年）。市場化テストと呼ばれる官民競争入札、民間競争入札を導入。
- 指定管理者制度（地方自治法の一部改正、2003年）。道路、下水道、河川、病院、学校などさまざまな施設が対象。管理者の選定に評価。

さまざまな評価を導入し、経験した地方自治

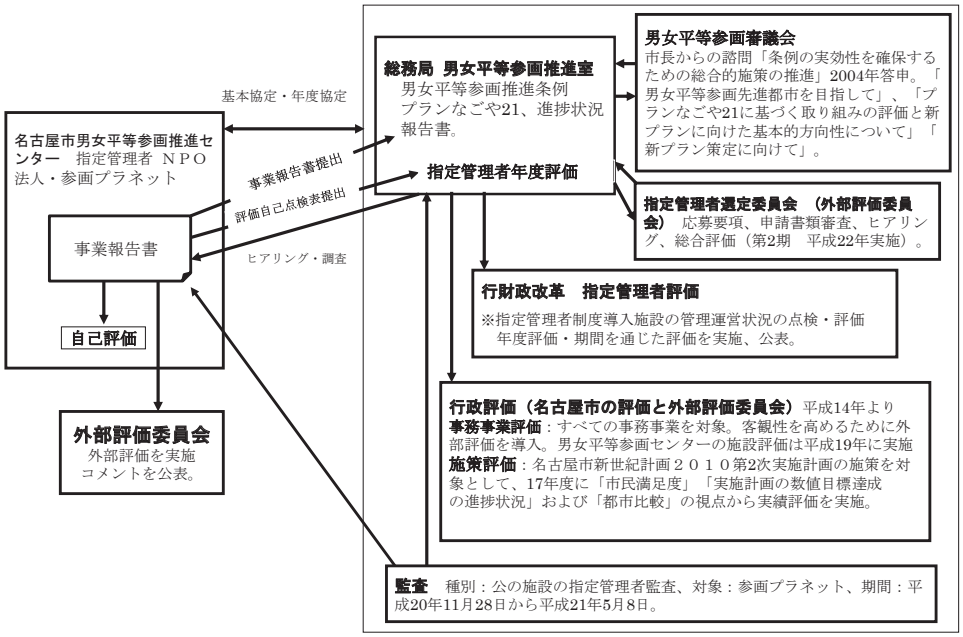
体には、大きな問題がある。すなわち、目的も考え方も違う「評価」が同時に存在し、2～3年ごとの人事異動でやってきた評価素人の担当者がその違いを十分認識しないまま実践している「評価の氾濫」である。氾濫を招いた直接の原因は、各課がそれぞれの業務の中で勝手に導入した評価体制にあるが、実際は上記の民間委託の枠組みの中で事業を受託した事業者が、一次評価をする場面が多い。その典型的な事例が男女共同参画関連事業である（【図1】を参照）。

さまざまな評価の配置を示したこの評価マップに見られるように、指定管理者に選ばれたNPOをたくさんの評価関連業務が待っている。【図1】の例では、まず①市や県が定めた計画の進捗状況報告を事業報告書の形で提出、②計画に基づいて実施される事業の自己評価、③事業報告に対する外部評価委員会向けの対応、④総務局男女平等参画室に向けて指定管理者に対する男女平等参画室からのヒアリング対応、⑤監査への対応である。

市側も⑥指定管理者年度評価の報告、⑦指定管理者選定委員会への評価報告、⑧行政評価（事務事業評価・施策評価）を行わなければならない。これら8種類の評価の方法を理解し、書式を考え、タイミングよく公表する必要があるが、実際には評価全体をコントロールする司令塔がないため、評価や調査、報告書の提出、監査を各課はNPOに対して「さみだれ式」に求めるのである。しかも、NPOが評価理論を勉強し、評価知識を活かして巧く対応したとしても、次期の指定管理者に選定されるとは限らない。指定管理者は8つの評価以外の、入札金額で選択されることが多いからである。

しかも、評価の氾濫と併行して、「定員管理」の名目で地方公務員定数がずっと削減されてきた事実を一般市民は無視してきた。総務省のホームページによると、ピークの1994年に328万2000人であった地方公務員総数は、2015年には273万8000人にまで削減されている。とくに評価に関わる総務や管理部門は、マンパワー不足が恒常的に進んでいる。同一人物が監査委員会事務局、公平委員会事務局、そして行政評価を兼

図１　評価マップ（名古屋市男女平等共同参画センターの例）



（出所）林泰子「NPO評価と行政評価の相剋」、日本評価学会春季第10回全国大会（2013年5月26日）、報告要旨収録、p.118を一部修正して引用。

務する事例は珍しくない。そして評価関連業務が増え続けているのである。残業を前提として業務を処理する中で疲労する、評価が「制度疲労」を招くという言いたい真意はここにある。

他方、公務員には人事異動があり、同じ部課の中でも別の業務担当への異動が多い。結果として評価担当者も2〜3年おきに異動になり、常に半ば素人が評価を担当するが、評価に馴れてきたと思う頃には再度異動である。新しく着任した担当者は前任者が置いていった資料を見ながら、にわか勉強の自己学習で評価業務を何とかこなす。反面、評価関連業務の理屈、理論は年々精緻化、複雑化をたどる。素人が業務処理のかたわら、この理屈や理論をフォローするのはかなり難しい。もちろん、公務員の採用試験には無い会計学や評価学の専門に習熟した公務員が現場に多いとは思えない。

しかし、21世紀になって地方自治体でとくに求められ続けてきたのがコスト削減であれば、目標管理型で業務活動を測定するのが一番無難である。行政活動に目標を付け、その達成度を測定し、確認された数字を見て予算の議論をする現実的方法である。青少年健全育成、男女共同参画、防災PRなどいろいろな分野に横断的に使えるし、特別に勉強をする必要もない。業績測定は、じつは現実的対応なのである。

5　新たな評価ー「地方創生」

ところで、1990年代から続いた過疎地域の「限界集落」化問題が広く認知された後、こんどは「壊死する地方都市」（『中央公論』2013年12月号）、「消滅する市町村」（『中央公論』2014年6月号）、「すべての町は救えない」（『中央公論』2014年7月号）と衝撃的なタイトルを特集にした雑誌が現れた。このままでは多くの集落が現状維持できなくなり消滅する、そのような危機感を煽ったのである。少子化・高齢化社会が来ているのに、中央府省も地方自治体も積極的な対応を先送りしてきた問題、その先にあるものを、この三つの特集はリアルに、そしてセンセーショナルに示したのである。

「地方創生」はこうした危機感に対する自民党政権の対応として生まれた。2014年9月、人口減や少子高齢化などに取り組む「まち・ひと・しごと創生本部」が設置されたが、これは内閣総理大臣を本部長、地方創生担当大臣と内閣官房長官を副本部長、ほかのすべての国務大臣を本部員とする組織である。その現実はどのようなものであろうか。

「地方創生」の場合、地方に提案を出させるスタイルであるため、その提案能力の有無が交付金獲得を左右する。まさにこの部分が、過疎地域にある小規模自治体を苦労させる。なぜなら、前述の事例のように、小規模自治体では職員数が少ないためにこうした提案を考える作業にエネルギーを割く余裕がないからである。そもそも、提案の中にはPDCAを巧くまわすためにKPIを入れなければならないが、KPIの選択もまた手間がかかる難問である。

政府が用意した政策、地方創生を具体的に語る「まち・ひと・しごと創生総合戦略」（2014年12月）は、従来の地域政策について五つの反省を示す。府省庁・制度ごとの「縦割り」構造、地域特性を考慮しない「全国一律」の手法、効果検証を伴わない「バラマキ」、地域に浸透しない「表面的」な施策、「短期的」な成果を求める施策、この五つである。この五つの反省にもとづき設定した原則が、「まち・ひと・しごとの創生に向けた政策5原則」である。

- 自立性。各施策が一過性の対症療法的なものにとどまらず、構造的な問題に対処し、地方公共団体・民間事業者・個人等の自立につながるようなものにする。
- 将来性。地方が自主的かつ主体的に、夢を持って前向きに取り組むことを支援する施策に重点を置く。
- 地域性。国による画一的手法や「縦割り」的な支援ではなく、各地域の実態に合った施策を支援する。各地域は客観的データに基づき実状分析や将来予測を行い、「都道府県まち・ひと・しごと創生総合戦略」および「市町村まち・ひと・しごと創生総合戦略」（「地方版総合戦略」）を策定し、この戦略に沿った施策を実施できる枠組みを整備

する。

- 直接性。ひとの移転・しごとの創出やまちづくりを直接的に支援する施策を集中的に実施する。地方自治体、住民代表、産業界・大学・金融機関・労働団体（産官学金労）の連携を促し、政策の効果をより高める工夫を行う。
- 結果重視。効果検証の仕組みを伴わないバラマキ型の施策は採用せず、明確なPDCAメカニズムの下に短期・中期の数値目標を設定し、政策効果を客観的な指標により検証し、必要な改善等を行う。

この5原則の成功の鍵になるのが、現場の効果の検証に使うPDCAサイクルだと言うのである。すなわち国も地方も地域の課題をデータに基づき抽出し、その解決に向けて計画（plan）を策定し、実施（do）する。この実施状況がうまく進捗しているかどうかを判断するために成果（アウトカム）に重要業績指標KPIをつける。このKPIに基づきチェック（check）し、そのチェックした結果を政策の現場にフィードバックして改善の行動（action）をとらせる、こういった仕組みがPDCAサイクルである。

ところで、KPIに代表される業績測定と呼ばれる手法は、一般に6段階に分かれる。

①ターゲットの選定。組織、政策・プログラムのいずれが測定対象なのか、その確認から始まる。政策・プログラムを対象にするときには、プログラムのロジック（インプット→活動→アウトプット→アウトカム→インパクト）が適正か見る必要がある。また組織を対象にするときには、戦略的プランニングで用いられる目標の「ツリー」が必要になる。つまりミッションを表明し、このミッションを戦略目標に読み替え、その戦略目標を現場が達成すべき具体的な目標に翻訳する。もちろん、この現場の目標については資源をどのように確保するのかを明らかにする。

②指標（indicators）の選定。投入するインプットの指標、インプットを使って産出したいアウトプットの指標、アウトプットによって形成される中間アウトカムの指標、最終アウトカムの指標、環境に及ぼす影響やインパクトの指標である。

③実効性確認。選定したターゲットの妥当性、指標の適切性などをチェックして、実際に有効なのかどうか、意味があるのかどうかを確認する。ただ、ターゲットや指標を地方自治体のトップが見ることは重要だが、指標達成を無理に強制すると問題が生じることがある。代表例が大阪府警のデータ改竄事件である。橋下徹元大阪府知事の『街頭犯罪ワースト1返上』方針が府警の改竄を誘発し、2014年7月31日、橋下氏が記者会見で謝罪した事件である。正しいデータを提出することは大事であるが、組織防衛が先立つと、その正しさを担保できなくなる事例である。

④データの収集。既存の統計データがある場合にはその活用で十分かどうかを事前確認、もしそれが十分でないときには調査・リサーチで収集する。この収集費用を負担できるかが問題になる。またデータの客観性の問題や、素人には難しいエコノメトリクスやサイコメトリクスで用いられるようなテクニカルな測定が可能かどうかも考える必要がある。

⑤分析。分析とはデータをよりメッセージ力の高い、見た人が即時に理解できるレベルのインフォメーションに変換するプロセスのことである。このプロセスが必要な理由は、生データそのままでは意味をなさず、役に立たないからである。一定の視点から整理して、役に立つ情報に加工する作業（＝分析）が必要になる。時間経過によって表示する、他の組織と比較する、場所ごと、対象集団別、地域別のデータ分析もあり、技術的分析やリスク分析もある。

⑥報告。誰が、どんなタイミングで、いかなる情報を求めているのかがこの段階での重要事項で、これを確認しないまま一般的な情報提供をしても、それはゴミでしかない。考慮すべき事項は以下である。報告を求めている人にとっていま使っている報告のフォーマット、票やシートは適切か。報告のタイミングをどうするのか。とくに、多くの地方自治体が使うシート記入方式の課題は、シートを見る人、必要とする人が誰なのか絞りきれないまま漫然と、年中行事として書き込んでいる実態である。

一般的には業績測定のプロセスを考えることは、この6段階をどのように処理し、進めていくか、そのプロセスの構築である。もちろんKPIもこのプロセスを前提にしている。

ところで、「まち・ひと・しごと創生総合戦略」は非常に良くできた戦略文書である。たとえば、付属文書としてアクションプラン（個別施策工程表）全82ページが添付され、その精巧かつ具体的な付属文書を見れば、普通の地方公務員であれば何を、どのように実施すれば良いか、簡単に分かる仕組みになっている（【表2】を参照）。もちろんKPI事例も数多く示されているため、地方自治体側もそれに倣ったKPIを設定すれば良い。KPIをめやすにしてPDCAサイクルをまわし成果を出すと考える思考がここにはある。

PDCAとKPIは生産現場の品質管理、QC活動に出自を持ち、トヨタの「カイゼン」のような場で実力を発揮する。したがって、現場から遠く離れたところにいるトップの執行幹部も、KPIを見れば現場で何が起きているのか数字で確認できる。

しかし、KPI方式には短所もある。「まち・ひと・しごと創生総合戦略」のように国が詳細かつ具体的なモデルをKPI付きで出すと、地方の小規模自治体はそれをモデルに総合戦略を作り、国が示したKPIに準じた指標を付けていく。とくに、その戦略の出来・不出来を審査（評価）されて、その結果で交付金がもらえるかどうかが決断されると、地方自治体ではモデルに一体化したい誘惑が強まる。一種の思考停止である。逆に、国はKPIを「コンプライアンス・チェック」項目に使いたい誘惑に駆られる。とくに戦略や計画が事業仕分けや行政事業レビューの対象になった場合、誘惑が強まる。国の考える成果指標をなぞって地方自治体が指標を設定し、それを「きちんと達成できています」と言えば廃止

表２　「まち・ひと・しごと創生総合戦略」付属文書の工程表の例

（２）地方への新しいひとの流れをつくる	（ア）地方移住の推進

(2)-(ア)-①　地方移住希望者への支援体制

●現在の課題

○東京都在住者の約４割（うち関東圏以外出身者の約５割）が地方への移住を検討又は今後検討したいと考えており、特に若年層や50代男性の移住に対する意識が高いとの結果が出ている。

○地方移住関連情報は、主として市町村がインターネットで提供しているが、精粗まちまちであり、体系的な情報提供体制が整備されておらず、キーワード検索ができない状況である。

○特定非営利活動法人や一部の都道府県がそれぞれ相談支援を行っているが、各都道府県の連携やしごとの相談との連携が希薄であり、一括した情報の提供や場所の認知がなされていないなど、移住を考える者の相談に一元的に対応できる体制が未整備である。

●必要な対応

○住まい、教育、医療などの生活面の情報だけでなく、農林漁業、企業等求人情報も加え、移住に必要な情報の一元的な提供システムを整備する。

○全国の地方受入組織と連携した移住関連情報の提供・相談支援を一元的に行う「全国移住促進センター（仮称）」を整備する。

○地方公共団体が実施する移住希望者に対する移住関連情報の提供や相談支援について、地方財政措置を創設する。

●短期・中長期の工程表

	緊急的取組	2015 年度	2016 年度以降（５年後まで）
取組内容	○移住に必要な情報の一元的な収集・提供体制の構築 ○移住についてワンストップ相談ができる「全国移住促進センター（仮称）」の開設 ○受入側となる地方の体制整備、連携の促進 ○移住促進に対する政府の取組を国民に広く周知	○一元的な情報提供システムの本格稼働 ○「全国移住促進センター（仮称）」の本格稼働 ○受入側となる地方の体制整備、連携の促進 ○地方公共団体に対する地方財政措置	○情報コンテンツの順次充実 ○「全国移住促進センター（仮称）」の活動と各都道府県が行う移住の相談支援事業との連携拡大
2020 年 KPI（成果目標）	○地方移住施策全体の推進を通じて 2020 年に東京圏からの転出４万人増加（2013 年比） ○2020 年に「全国移住促進センター（仮称）」のあっせん件数 11,000 件		

（出所）内閣府「まち・ひと・しごと創生総合戦略」（20014年12月27日）、付属文書「アクションプラン（個別施策工程表）、38ページ。本表は官邸ホームページ記載文章より抜粋して作成。www.kantei.go.jp/jp/singi/sousei/pdf/20141227siryou6.pdf（2016年5月26日最終閲覧）。

対象にならないからである。PDCAとKPIにはこのように、思考レベルでの中央統制の危惧がつきまとう。

それでは市民自治の視点でKPIを考え、PDCAサイクルをまわす方法があるのだろうか。答えはイエスである。

ただしその大前提で必要なのは、KPIも、PDCAサイクルも、地方版戦略の策定も、地方公務員がすべて管理することをしない、これである。地域住民を集め、熟議してもらうことは誰でも理解している。ただし、その会合のロジスティックスを地方公務員が担当し、会議をリードしていくのは無理である。無理だという理由は、市民参加の場面である。一般の市民はウィークディの昼は集まれないため夜6時半から開始する。それで足りない場合は、土曜日や日曜日の午後に行うことになる。それには公務員を疲労困憊させてしまう。この場合には地元NPOを活用する方法がある。ただし、地元のNPOが評価を分かっているかどうかが重要なポイントで、評価が分からないNPOに任せると議論は混乱し、結論が出ない。「まち・ひと・しごと創生総合戦略」では地域市民とNPOを重要なパートナーとして位置づけているが、こうした文脈でNPOの育成、キャパシティ・ビルディング、あるいはエンパワーメントが地方創生の必須条件になってくる。

さて、これまでの議論で既に明らかなように、政策評価、行政評価、そして氾濫した各種の評価では行政内部のマネジメント、内部管理志向がきわめて強かった。しかし、内部管理、内部のマネジメントに拘泥する思考は、地域社会の改革を目指す上で限界に来ている。

ここに市民の参加型評価が求められ、その前提として行政と市民とのコラボレーションによる協働型評価が試みられるべきであり、しかしこのためには市民に対するエンパワーメントが不可欠という、どこかで耳にするデジャヴのストーリーが見えてくる。さらに新しく18才選挙権も登場する。将来の地域の担い手である若者をどのように覚醒させるか、その考えの道筋は評価にある。

まとめ

政策評価は選挙を経て首長になった政治エリートが、国全体の政治課題であった地方分権改革に積極的に取り組むツールであった。行政評価は、財政問題を抱える地方自治体における経営健全化をめざす行政内部の管理改善運動であった。その後に出てきた評価の氾濫状態は、NPM型改革の流れをくむ市場化ツールに伴う評価を受け入れた結果として出てきた。この三つの評価ブームは、地方の公務員が勉強しなければならないことが多く、とくに政策評価と行政評価は公務員・行政の意識改革の面も強調されたため、日本各地で研修、勉強会が開催された。全庁的なカイゼンのイメージである。

NPM型の諸評価は平成の大合併、市町村合併を経て状況が変わり、総務系の自治体職員以外がイニシアチブをとる場面は減った。福祉でも、男女共同参画でも専門職の評価はあるけれども、そうした評価以外に業績測定やランキングや評定が行われる別次元の評価であった。いずれにしても、これらは現場に力点があり、地方自治体の第一線の現場職員から総務系職員に上がっていく、ボトムアップ的な性格が強かったのである。

これに対して、地方創生のPDCAサイクルとKPIは、全く違う。明らかにトップダウン、それも国（中央）から指示された具体的方法、細部まで決められた進め方で動いている。まさにグローバル大企業がトップダウンで戦略を決め、その進捗状況を鳥瞰するためにKPIを使う、この方式によく似ているのである。ここでの評価、測定は中央のコントロール手段であると言い換えられる。重視されるのは、現場がどれだけ戦略に一体化しているか、トップがどこまで戦略の実施体制、すなわち現場をコントロールできるか（controllability）であり、それがアカウンタビリティの重要な要素になってくる。戦略が失敗する場合にはトップがその責任をとる仕組みも求められ、もし責任をとらないのであれば、戦略は無責任の体系になる。そのうえ、地方自治体で2～3年ごとに繰り返される人事異動は無責任の体系を招く危うさを内在している。KPI型

の業績測定のマイナス・イメージはここにある。無責任の体系の中で、現場の地方自治体は交付金につられて洗脳されるか、何も考えず思考停止に陥る恐れが強い。

欧米での評価の研究と実務の潮流を時代にしたがって、4つに分ける研究があり（Vedung 2010）、本稿はこれを参考にした。すなわち、①1950年代から1960年代半ばまでのサイエンス主導のラジカルな合理主義、エンジニアリング重視の波（the Science-driven Wave）、②1970年代半ばから対話志向の波（the Dialogue-oriented Wave）、③1978年から79年頃に揺籃期を迎え、その後各地に拡散した新自由主義の波（the Neo-Liberal Wave）、④1995年頃から始まったエビデンスを重視する評価の波、の四つである。評価を専門として研究してきた立場からすれば、この4つの波に分類する主張には納得できる部分が多い。もちろんこの波のそれぞれにしたがって、評価手法群におけるスキルやツールの主役が変わり、あるいは洗練を遂げてきたことは評価の専門家は承知しているし、この4つがそれぞれ一時の流行に終わるのではなく、重層的に蓄積していることもまた理解している。国際的に見れば、評価は多様化しているのである。

しかし、日本の場合、問題が二つある。一つは評価は普及・拡散・重層化したのではなく、一点に収斂してきた現実である。収斂の先は目標管理型の業績測定である。評価素人が扱うことを前提にすれば、分かり易く、数少ないスタッフでも担当できる省エネ的な方法が求められる。また、評価結果を予算に使うため評価は数字を求める「測定」「指標」の方向に集中する。逆に言えば多様性を排してきた結果がこの目標管理型業績測定であり、それをPDCAサイクル、KPIとして使っている。もちろん、このPDCAサイクルは、経済協力開発機構（OECD）の開発援助委員会（DAC）や世界銀行の評価関係者の会合で日本からの出席者が言及しても、他国の関係者は理解できない和製英語である。

二つめの問題は、こうした評価・測定の波が欧米では60年の長期にわたって行われ、そのため実践に理論が追いつき、その理論を実践が吸収する余裕があった。しかし、日本ではたかだか20年であり、理論と実践の交流、相互発展にとっては短すぎたのである。

国際的潮流とは違った方向に進化する日本国内の行政評価は、皮肉な表現であるが「評価のガラパゴス化」と言うべきではなかろうか。

参考文献

窪田好男（1998）「三重県の事務事業評価システム　評価を通じた政策改善、あるいは評価システムによる『行政革命』」、日本公共政策学会年報

山谷清志（2012）『政策評価』、ミネルヴァ書房

山谷清志･編著（2010）『公共部門の評価と管理』、晃洋書房

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

Osborne, David and Gaebler, Ted(1992). *Reinventing Government*, A Plame Book.（高地高司・訳、野村隆・監修（1995）『行政革命』日本能率協会マネジメントセンター）

Vedung, Evert(2010). ‘Four Waves of Evaluation Diffusion,’ *Evaluation*, 16(3), 263-277.

（2016.6.7 受理）

Many faces of evaluation in Japanese local governments －20 years of Decentralization and the Financial crisis－

Kiyoshi Yamaya

Doshisha University
kyamaya@mail.doshisha.ac.jp

Abstract

Since around mid-1990s, Japanese local governments, for example Mie prefecture, Iwate prefecture, Akita prefecture and Hokkaido prefecture, have tried to use policy evaluation, project analysis. In addition to evaluation and analysis, many local governments used the management review, management-by-objective type organizational activity measurement, performance measurement. The methods and skills of evaluations and measurements are called the ‘administrative evaluation’ system. In the 2014, central government (Cabinet Office) has used the PDCA (plan-do-check-act) cycle and KPI (key performance indicator) system for grants-in aid and subsidies for local governments. Many types of evaluations and measurements have converged on the PDCA and KPI. This convergence is unaccountable problem in the Japanese evaluation theory.

Keywords

policy evaluation, performance measurement, ‘administrative evaluation’, PDCA/KPI

【研究ノート】

ネットワークによる公共サービスのマネジメント・ガバナンスと業績評価
－欧米の先行研究に見る動向と今後の研究課題－

荒川　潤
慶應義塾大学大学院
政策・メディア研究科後期博士課程
arakawa525jun@gmail.com

玉村　雅敏
慶應義塾大学総合政策学部 教授

要　約

複雑な政策課題や多様な政策ニーズなどに対応するために、複数の政府機関間や政府機関と民間組織（民間企業や非営利団体など）との間でネットワークが形成されるケースがある。本稿は、そのようなネットワークのマネジメント、ガバナンスの在り方について、及びそこでの業績評価の在り方について、欧米における主要な先行研究をレビューしたものである。

ネットワークの形成により、非定型業務や多様性のある業務に柔軟に対応できるようになるなどのメリットがある一方で、多様な利害や関心を有する多様な主体から構成されるネットワークについて、それを円滑にマネジメントすることや、適切なガバナンスを確保することには、単一組織の場合よりも、より一層の複雑さや困難が伴いがちである。例えば、複眼的に森と木とを同時にみる、複数のシステムを同時にまわす、複層的に取組む、など複数のことに同時に取組むことが求められる。また共有・協働と委譲・裁量など、方向性の異なるプロセスにも同時に取組む必要がある。そして、構成員相互の信頼関係や利害調整の重要性が強調されるのも特徴的である。

今後、わが国においても理論・実証の両面からの研究がさらに蓄積され、それらが「PPPの失敗」を含めた実務的な諸課題に対する処方箋となることが期待される。

キーワード

ネットワーク、マネジメント、ガバナンス、業績評価、業績情報

1．はじめに

政府機関間もしくは政府機関と民間事業者や非営利組織などの複数の主体によるネットワークが形成されて、それにより公共サービスが提供され、政策課題の解決が目指されるケースがある。象徴的には、大災害時の救助・復旧・復興などにおけるネットワークが想定しやすい事例であろう。また非常時に限らずとも、平時の通常の公共サービス提供においても、ネットワークが形成されることも少なくない。

そのようなネットワークによる公共サービスでは、政府単体や単なる外部委託を通じてサービスが提供される場合とは異なり、参加する複数主体間の密接な連携により一つの価値が創造され提供されることになる。この特徴は、公共サービスの

マネジメント及びガバナンスの両面において、またそれらにおける業績評価の在り方や役割について、単独主体や外部委託による公共サービスとは異なるものを求めることとなる。

具体的に、ネットワークによる公共サービスの場合には、ネットワークが総体として成果を出しかつ説明責任を果たす必要があると同時に、参加する各主体もそれぞれの立場から成果を上げて説明責任を果たすことが求められる。従って、そのマネジメントやガバナンスも、またそれを支える業績評価も、このネットワークならではの要請に的確に応えていく必要がある。

この点については、特に欧米諸国において、理論・実証双方の観点からの多くの先行研究が存在している。本稿は、これら海外の先行研究の動向を把握して概観するとともに、わが国における今後の研究課題を整理することを目的とする。

なお本稿で「ネットワーク」とは、Osborne（2010a）、Kettl（2009b）、Mandell and Keast（2008a）などを参考にして、複数の多様な主体が参加して、それぞれが機能・責任を担いつつ（リスクを分担しつつ）相互に依存しながら、共通の公的目的を追求する業務形態、すなわちそれら主体の連携により共通の「公の価値」を創り、それを市民や産業界などに提供することにより政策課題の解決に貢献する業務形態を指している。具体的には、コンセッション（施設の所有権は官が保有し、その運営権は民が保有する方式）やPFI（民間の資金やノウハウを活用して公共サービスを提供する方式）などが、その例として想定される。[1-2]

また「マネジメント」とは、ある組織体における成果創出（目標達成）や法令遵守などの実現に向けた執行行為を、更に「ガバナンス」とは、ある組織体における成果創出（目標達成）や法令遵守などの特定の価値の安定した実現を確実にするために、利害を有する特定の立場から、意思決定、舵取り、監視、（情報提供や説明の）請求、（必要に応じた）修正の請求などを行うことを、それぞれ指している。

また特に断っていない場合には、政府機関間ネットワークと官民ネットワーク（政府機関と民間企業や非営利組織などの民間主体とのネットワーク）とを、区別せずに議論する。

2.　ネットワーク化の進展

（1）ネットワークの形成

政府による公共サービスの提供に際して、ネットワークの組成が促進される条件（ネットワークが必要とされる局面）として、Moore（2009, p213）は、①政府機関が成果を出せないことが目立ち危機的な状況にある場合、②対応すべき政策課題と対応する政府機関との間にミスマッチがある場合、③解決すべき政策課題の特性に照らして単独政府機関による対応よりもネットワークによる対応がより適切な場合、の3ケースがあると分析する。③は、例えば環境問題のように、その原因がある場所と結果（汚染等）が発生する場所とが遠隔となりうる場合や、例えば犯罪対策・肥満対策など小さな政策課題の集合体として大きな政策課題があるケースのように多段階の多様な対応が同時に求められる場合（国・自治体・コミュニティなどでそれぞれ対応が必要な場合）、などが想定されている（p216-7）。

（2）ネットワークの種類

政府機関が構築するネットワークの種類としては、①協力ネットワーク（Co-operative Networks、独立した主体が必要に応じて情報や専門性を交換するもの）、②協同ネットワーク（Co-ordinative Networks、独立した主体が個々の取組みを統合させて協同でより良いサービス提供を目指すこと）、③提携ネットワーク（Collaborative Networks、相互依存する主体が各主体単独では解決しえない課題の解決のために提携すること）の3種類があると、Mandell and Keast（2008a, p690）は指摘する。

上記のように、本稿では、主として③を想定して議論を展開する。

（3）政策課題への対応戦略：ネットワークはいつ有効か

公共サービスの提供に際してネットワーク化が急速に進展していることと、どの政策課題をネットワークによって解決するのかということとは別の問題である。Goldsmith and Eggers（2004, p51）は、問題の特質などにより、ネットワークモデルにふさわしい場合と、ネットワークではない旧来

の官僚的組織モデルの方がふさわしい場合の双方があると分析する（表1参照）。同様にKettl（2009a, p99）も、政府の対応方法には、旧来の階層組織を用いる場合と複雑なネットワークを用いる場合があるとしている。全体として、非定型性のある政策課題の場合、課題への対応に柔軟性や多様性などが求められる場合、民間の能力活用が見込める場合や急速な技術革新があるような場合に、ネットワークを用いることが想定されている。Kettl（2009a）では、米国の公的医療保険制度であるメディケア・メディケイドや、大災害ハリケーン・カトリナなどへの対応、NASA におけるロケット開発、などのネットワーク事例を用いて議論が進められている。

3. ネットワークのマネジメント、ガバナンス

（1）ネットワーク・マネジメントの特質

O'Leary and Bingham（2009, p264-5）は、ネットワークで他機関と連携しつつ業務に携わるマネージャー（collaborative managers）は、自らの組織とネットワークの双方で同時に仕事をするため、様々な「パラドックス」に直面すると分

表１　政府の選択を決める要素

ネットワークモデルがふさわしい要素	官僚的組織モデルがふさわしい要素
○ 柔軟性が必要な場合	○ 安定性が必要な場合
○ 顧客に対して異なった対応をおこなうことが必要な場合	○ 画一的で、ルールに従った対応が必要な場合
○ 多様なスキルが必要な場合	○ 単一の専門性が必要な場合
○ 多くの潜在民間プレーヤーが見込める場合	○ 政府が支配的な供給者である場合
○ 望ましい結果・成果が明確な場合	○ 成果が曖昧な場合
○ 民間がスキルのギャップを埋めうる場合	○ 政府が必要な経験を有する場合
○ 民間の資産の活用が重要な場合	○ 政府以外の能力が重要ではない場合
○ パートナーが能力や信用力を有する場合	○ 政府がその地域の市民との経験がある場合
○ 複数のサービスが同じ利用者に接する場合	○ サービスが比較的独立して存在する場合
○ 第三者が政府よりも安いコストでサービスを提供できる場合	○ 政府によるサービスの方がコストが安い場合
○ 急速な技術革新がある場合	○ 技術革新に影響されない場合
○ 複数レベルの政府によりサービスが提供されている場合	○ 単一レベルの政府によりサービスが提供されている場合
○ 複数の機関が同様の機能を用いる場合	○ 単一の機関が同様の機能を用いる場合

（出所）Goldsmith and Eggers（2004）

表２　ネットワーク・マネジメントの課題

課　題	概　要
ネットワークの目的に対するバラバラの関与	◆ネットワーク成果と自身の機関の利害の重なり具合がマチマチであり、それが関与の程度の差を産み出す。
文化の衝突	◆（多様性は強みでもあると共に）協働を困難にしかねない要素にもなる。
自主性の喪失	◆ネットワークの決定が自身の機関の利害を反映しない懸念がある。
コーディネーションの疲れと費用	◆意思決定と実施に時間と労力を要する。 ◆適切なネットワーク・ガバナンスの導入が不可欠である。
説明責任の低下	◆成功・失敗を特定の機関に帰すことができず、説明責任が拡散する。「フリーライダー」を生み出す懸念がある。
マネジメントの複雑さ	◆自身の組織とネットワークの二股をかける必要性がある。 ◆ネットワークの目的にかかる参加者間の緊張・紛争がある。

（出所）Provan and Lemaire (2012, p642) の該当部分の記述を筆者が要約

析している。例えば、ネットワーク全体で共有する目標と自らの組織に固有な目標とを同時に追求することなど、同時に「複数の顔」を持って仕事をすることが求められる。

同様に、多くの機関が共通の目的に向けて協働して業務遂行する際に発生するマネジメント上の課題（challenges）として以下が指摘されており、ネットワーク・マネジメントの難易度が高いことが分かる（表2参照、Provan and Lemaire 2012, p642）。つまりこのようなネットワークでは、参加主体の価値観や利害は多様であり、そのような多様な主体がネットワークの共通の目的の実現に向けて共同作業する際には、煩雑な調整業務が必要になると共に、時に文化の衝突も避けられない。また、必ずしも自らの利害がネットワーク全体の意思決定に反映される訳ではないことから、各主体としての成果の追求や利害関係者への説明責任にも制約が生じることとなる。

(2) ネットワーク・マネジメントの類型

そして Milward and Provan（2006, p18）は、ネットワークのマネージャーが果たすべき役割は、「ネットワーク全体のマネジメント（Management of Network）」及び、「ネットワークに参加している各機関のマネジメント（Management in Networks）」の双方があるとした上で、それぞれについて、説明責任や正統性の確保、利害対立への予防と対応、ガバナンスの確保、ネットワークへの関与、の観点からの対応が必要である旨を体系的に整理している。

なお、ネットワークのマネジメントに関しては、他にも、ネットワークにおけるマネジメントの「チャンネル」と「機能（role）」に着目して類型化する研究（Voets 2014）などもある。

(3) ネットワーク・マネジメントの担い手

次にネットワーク・マネジメントの担い手について、Milward and Provan（2006, p22）は、大別して、①ネットワーク参加者相互によるもの、②ネットワーク内のリード（lead）機関によるもの、③別途雇用された第三者のネットワーク管理者によるもの、という3形態があるとする。同様に、これとは少し観点が異なる部分もあるが、Goldsmith and Eggers（2004, p75）は、官民ネットワークを束ねる「統合役」の重要性を指摘した上で、その「統合役」として、①政府が自らその役割を果たす、②主契約者に統合役の機能を委託する、③第三者の統合役を別途雇用する、という3つの選択肢があることを示している。

なお、ネットワークのマネージャーは、必ずしも一人ではないかもしれない。複数のマネージャーによる集合的な活動の結果として、求められる機能が果たされることもある。（McGuire and Agranoff 2014, p152）

(4) ネットワークのガバナンス

次にネットワークのガバナンスについては、例えば説明責任について、複数の枠組みを同時に機能させることの重要性が指摘されている。政府間ネットワーク・官民ネットワークの何れの場合も、多様な主体がそれぞれ多様な利害関係者（出資者・融資者、サービス利用者、市民、議会、他レベルの政府など）を抱えており、それらに対する説明責任を果たしていくためには、複数の枠組みが必要になる、との考え方である。具体的に Kettl（2009a, p223）は、ネットワーク化が進展する時代を多様な主体による活用型ガバナンス（Leveraged Governance）の時代と位置付けて、「複数の説明責任システムを、機動的・動態的・効果的なものにする」ことの重要性を指摘している（表3参照）。また Koliba, Meek and Zia（2013, p36, p243）は、「民主的な」「市場における」「行政的な」説明責任という3種類の説明責任を包括した枠組みを提唱している。公共サービスの特性によって、これら3種類の「混合的」な枠組みが構築される。

4. ネットワークの業績評価と業績マネジメント・システム

(1) ネットワークの業績評価の特質

本節では、ネットワークによる公共サービスに対する業績評価の特質を、従来型の公共サービ

表３　ガバナンス・モデルの進化

	古い公共経営	NPM	多様な主体による活用型ガバナンス
時代	工業化時代	情報化時代	ネットワーク化時代
説明責任の戦略	階層	結果（アウトプット）	共通の成果に対する各々の貢献
チャレンジ	非階層的な戦略・戦術の拡大	管理職は自ら責任を有していない成果を制御できない（が説明責任は持たされる）	複数の説明責任システムを、機動的・動態的・効果的なものにする

（出所）Kettl（2009a）。下線は筆者。

スにおける業績評価との比較（異同）の観点から簡単に整理する。

まず業績評価を実施する目的そのものは、ネットワークによる公共サービスも従来型の公共サービスも、どちらもマネジメント（進捗管理、執行の改善、資源配分、など）とガバナンス（説明責任など）に活かすという点で、同一であると考えられる。

その一方で、Koppenjan（2008）、Moore（2009）、Popp et al.（2014）などを参考にすると、ネットワークの業績評価は、例えば以下のような特色（必要性や困難さ）を持つ点で、従来型とは異なっていると考えられる。

・**複層性**：ネットワーク全体、ネットワークの各部門（例えば空港コンセッションにおける航空系業務や非航空系業務など）、参加主体、地域コミュニティなど、複数の層にて評価を実施する必要性がある点
・**双方向性**：従来型のような官による民のモニタリングのみならず、相互依存関係の下では、その逆も必要となりうる点
・**（業績情報やタイミングの）多様性**：ネットワーク全体の成果や説明責任に加えて、各参加主体がそれぞれにとっての成果を確認でき、かつ説明責任を果たしうるために、多様な業績情報が多様なタイミングで必要となる点（同時に、業績情報の過多によるマネジメント・ガバナンスの機能不全がより大規模に生じるリスクがある点）
・**判断の困難さ①**：業績結果についての判断が、利害が異なる参加主体によって大きく異なるリスクがある点（例えば、目標達成率80%をどう判断するか、など）
・**判断の困難さ②**：例えばネットワーク全体の業績は不変ながらも、その部門や小部門（例として、空港コンセッションの非航空業務における「有料駐車場管理」「商業施設管理」など）では顕著な業績が上がっている場合に、ネットワーク事業の成果をどのように判断するか難しい点

このようにネットワークでは、全体的に業績評価の複雑性が増していると言えよう。

（2）ネットワークの業績評価の視点

次に業績評価の視点について検討する。ネットワークによる公共サービスには、複数の視点からの業績評価が求められる。Provan and Milward（2001, p416）は、「コミュニティ」「ネットワーク」「組織・参加者」の3層における複層的な評価が必要になると指摘する（表4参照）。すなわち、ネットワークとして公共サービスの提供について成果をあげられたか否か、それが公共サービスの対象であるコミュニティ（社会）に対しても良い影響を及ぼしたか否か、ネットワークに参加している組織にとっても自らの利害に照らして良い成果があがっているか、の3層からの評価である。このProvan and Milward（2001）にて示されている枠組みは、現時点においても、ネットワークにおける業績評価の枠組みの原型として多くの研究者に引用されており、また新たに提案される枠組みの基礎となっている。

この内、組織・参加者を更に、「組織」と「個人」（individual）という2つに分割するという考え方もある（Popp et al. 2014, p79）。この中の「個人」については、①ネットワーク内で参加組織を

表４　ネットワークにおける評価の階層

評価の対象層	主たる利害関係者	「有効性」の基準
コミュニティ	一般市民 政治家 規制当局 出資者、など	・コミュニティへの費用 ・社会資本の構築 ・政策課題が解決されつつあるという社会の認識 ・政策課題の出現状況における変化 ・社会厚生の向上
ネットワーク	主たる出資者・規制当局 ネットワーク運営主体 ネットワーク参加組織	・ネットワーク参加者増 ・提供サービスの拡充 ・重複サービスの減少 ・ネットワークの強固さ ・提供サービスの統合・提携 ・ネットワークの維持費用減 ・ネットワーク目標への参加者の関与
組織・参加者	参加組織の経営者・役員会 組織の職員 組織固有の利害関係者	・組織の維持 ・自己の正統性の強化 ・資源獲得 ・サービス提供費用 ・サービスへのアクセス ・利害関係者への成果提供 ・ネットワーク間、参加者間の紛争の少なさ

（出所）Provan and Milsward（2001）

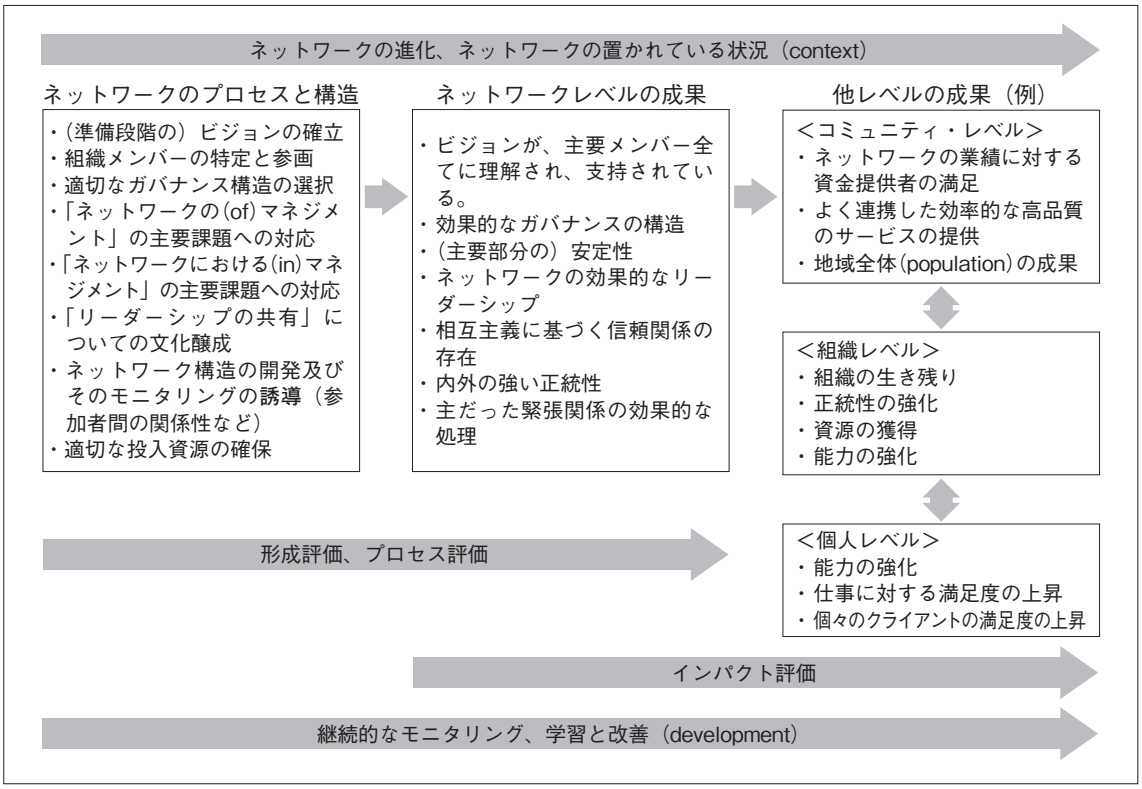
代表して活動する個人と、②ネットワークが提供するサービスを享受する個人という、立場の全く異なる2種類の「個人」が想定されている。

この他にも、ネットワークの業績評価を「生産」「プロセス」「体制（regime）」という視点から行う枠組みを提唱するVoets, Van Dooren and De Rynck（2008）や、「環境面」「組織面」「執行面」との局面から評価する枠組みを提唱するMandell and Keast（2008b）などの研究がある。また近年の研究として、Cepiku（2014, p178）は、上記Provan and Milward（2001）を含む既存研究を拡張・統合して、「ネットワーク業績の動態的多次元モデル」（Dynamic Multidimensional Model of Network Performance）を提唱している。ここでCepikuは、ネットワークの有効性の決定要因と中間成果（ネットワーク・マネジメントのアウトカム）、最終成果を構成要素として統合した一つのモデル体系を構築するとともに、それぞれの構成要素の評価基準（assessment criteria）を提示している。ここでは、ネットワークの有効性の決定要因として、内部資源、外部資源、外部要因、ネットワーク・マネジメントの4つが位置づけられている。そして中間成果（ネットワーク・マネジメントのアウトカム）は、ネットワーク内の相互作用の質を示すものとして位置づけられ、例えば参加主体間の相互作用（意思疎通など）の頻度や利害衝突の解決状況などが評価基準とされている。最終成果は、コミュニティ・レベルの業績と組織レベル（参加主体レベル）の業績から構成されており、前者では有効性・公平性・説明責任などが評価基準とされ、後者ではネットワークに参加しなかった場合の業績との比較、利害関係者とのより良い関係、などが評価基準となっている。

（3）成果の波及プロセス

成果の波及プロセスについてPopp et al.（2014 , p81）は、先行研究の体系的なレビューを踏まえて、ネットワークにおいて参加主体（コミュニティ、ネットワーク、組織、個人）にどのように成果が発生するのかを、ロジックモデル的に整理し

図１　ネットワークの評価に資するアクション・モデル



（出所）Popp et al.（2014, p81）

ている（図1）。[3]

それによると、ネットワーク全体の成果が出て初めて、参加組織や個人、そしてコミュニティへの成果が発生するとしており、この点がネットワークの特性であるとしている。また業績評価も、このような特性を十分に踏まえたものとすべきとしている。

しかし筆者ら自身が、このモデルは単純化して直線的なモデルとなっているものの、実際にはこれらのプロセスは反復（iterative）するものだと指摘するように、直線的に単純に成果が生み出されるというモデルではない。またさらに、後述するようにネットワークへの参加組織・個人が自分自身のネットワークでの業績を業績評価により確認できること（こそ）が、ネットワーク参加への関心・意欲を高めることにつながるとする考え方（Moore 2009、Goldsmith and Eggers 2004など）もあり、その視点からは、参加組織・個人の実績がネットワーク全体の成果につながるという流れ（図1で言えば、ネットワークレベルの成果から他レベルの成果に向けた右向きの矢印に加えて、反対向きの右から左への矢印）もあるものと思われる。

（4）プロセスの評価

Popp et al.（2014, p78）は、ネットワークの評価に関する近年の研究は、（成果を上げうる）ネットワークの構造についての研究が多く、プロセスに着目したものは少ないことに言及した上で、望ましい成果を導き出せるプロセスの在り方を把握することの重要性を指摘する。「どのように成果が達成されたのかを検討することにより、ネットワーク自体の健全性についての大事な情報を得ることができる。関係性の評価や、ネットワーク内に望ましい文化が醸成され維持されているのかついての評価などもそれに含まれる。」（p78）そして、ネットワークの構造とプロセスの双方につ

いての情報があれば、より良い状態に向けた修正が可能となるとしている。

この指摘は、上記（1）にて言及したネットワークの業績評価の特色（判断の困難さ②）に呼応した具体的な対応策を述べているものと考えられる。

（5）評価の指標

評価指標についてPoister（2010, p122）は、ネットワークの利害関係者の多様な関心を踏まえて、それらに適切に適合する複数の評価指標のセットが必要であり、場合によっては、重複する指標を複数の利害関係者に用意することもありうると指摘する。

この指摘も、ネットワークの業績評価の課題と実態を踏まえた、現実的な対応策を述べているものと考えられる。上記（1）で言及した特質を有するネットワークの複雑な業績評価を全体的に如何にしっかりまわしていくか、という点は、ネットワークの業績評価における主要な課題の一つである。Kettl（2009a, p265）が、ネットワークにおける「場所ごとの業績指標（Place-based Performance Measures）」を開発して徹底することの重要性を強調している点は、Poisterのこの指摘とも関連して、そのような課題に対する現時点での現実的な対応策と考えられる。ネットワークの施策体系に従った整然とした体系的な評価というよりは、ネットワークの各現場で成果をあげていくために業績評価が活用されることを主眼とした対応策と思われる。

5. ネットワークの業績マネジメント（業績マネジメント・システム）

（1）業績マネジメントの研究の類型

Koliba（2014, p90-98）は、ネットワークの業績マネジメントに関する先行研究のレビューを通じて、大別して3種類の研究があることを体系的に示している。そして、実際のネットワークのマネジメント実務は、これら3種を包含するとしている。

Kolibaによれば、まず第1のグループは特定の政策分野にて業績を産み出すための最適なネットワーク構造と機能を特定することを目的とする研究である。代表的な研究として、Provan and Milward（1995）、Meier and O'Toole（2003）があり、ネットワークの構造・マネジメント・業績の関係に焦点を当てた研究がなされた。またその後の研

表５　ネットワークの有効性の決定要因

カテゴリー	決定要因
ネットワークの構造 （structure）	・外部からのコントロール（外部への依存） ・ネットワーク統合のメカニズムとツール ・ネットワークの規模 ・承認（formalization）･･･規則、手続き、会義体など ・説明責任 ・ネットワーク内部の安定性
ネットワークを 機能させること （functioning）	・不安定性の防止（buffering）、安定性の促進（nurturing） ・ネットワーク・プロセスの舵取り（steering） ・伝統的なマネジメント業務（ネットワーク全体、各参加者） ・包括的なネットワーク強化・推進策
ネットワークの環境 （contextual）	・システムの安定性 ・資源の豊富さ（munificence） ・団結（cohesion）及びコミュニティからの支援
（結果としての） ネットワークの有効性	・顧客段階の有効性 ・コミュニティにおける有効性 ・設定した目標を達成するネットワークの能力 ・ネットワークの安定性・成長可能性・イノベーション・変化

（出所）Turrini et al.（2010）を筆者が表に整理した。

究であるTurrini et al.（2010, p533）は、ネットワークの構造のみならず、ネットワークを機能させることやその環境もネットワークの有効性を決定づける重要な要因であると分析している（表5参照）。すなわちネットワークのマネージャーによるネットワークの業務プロセスの柔軟かつ動態的なマネジメントや安定性の確保についての努力などもネットワークの有効性を左右することになる。

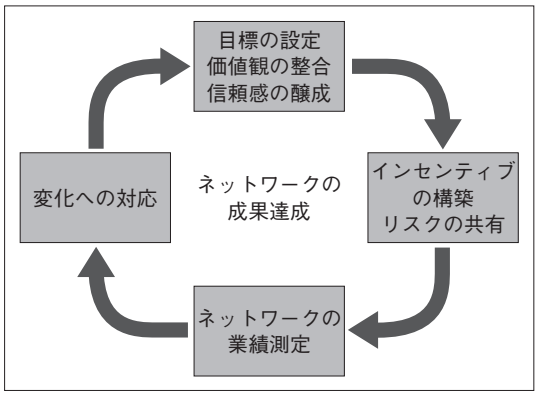
第2のグループは、業績マネジメント・システムの構築を通じネットワークが如何に業績をマネジメントするかの研究である。ここではネットワークにて業績情報が如何に収集され意思決定等に活用されるのかに焦点が当てられる。またネットワーク全体が、複数のサブシステムの集合体として位置づけられ、個々のサブシステムでの業績情報の活用が議論される。Kolibaによると、このグループの代表的な研究は、Frederickson and Frederickson（2006）、Koliba, Campbell and Zia（2011）である。前者はネットワークの業績評価（業績情報の収集と活用）に影響を及ぼすパラメーターに着目し、後者は一つのネットワークの中に複数の評価システム（サブシステム）が機能していることに着目する。

そして第3のグループは、ネットワークの業績マネジメント・システムとその成果とを結びつけることを試みる学習・知識の移転についての研究である。Kolibaによると代表的なものとして、業績マネジメント・システムを組織の学習とナレッジ・マネジメントを促進する枠組みとして位置づける研究（Moynihan 2008）や、成功するネットワークの特徴を分析する研究（Agranoff 2007）などがある。

（2）ネットワークにて成果を生み出すマネジメントの枠組み（第2のグループの例）

第2のグループのもう1つの例として、ネットワークが成果を生み出して、説明責任を果たすためのマネジメントの枠組みを提唱する研究がある。Goldsmith and Eggersは、ネットワークが成果を出すためには、「目標の設定」「価値観の整合」「信頼感の構築」「インセンティブの構築」「リスクの共有」「ネットワークの業績測定」「変化への対応」という7領域をサイクルとする包括的な枠組みが必要であると指摘する（図2参照、2004, p125）

図2　成果を産み出すネットワークの枠組み（全体イメージ）



（出所）Goldsmith and Eggers（2004, p125）

まず、個々の参加者の利害とネットワーク全体の目的とが整合するように調整した上で、適切な業績指標を用いてネットワーク全体の目標を設定する。次に、参加者間の価値観の調整と信頼関係の構築のために十分な時間を確保する。そして、ネットワークのガバナンスを効かせるにはインセンティブが重要であるので、業績に応じた報酬が支払われ参加者がしっかりと役割を果たすような仕組みを構築する。更に参加者間でのリスク分担が適切になされ、参加者により全体としてリスクが共有されている状態を構築することもガバナンスの観点で重要である。そして業績測定では、ネットワーク全体の状況と各参加者の機能状況の双方が確認されることが重要となる。最後に、動態的に変化する内外の状況に合わせて、参加者間の関係や契約条件を柔軟に変化させていくことが求められる。

またKettl（2009a, p264）も、「政府が成功するためには、ネットワークをつくったり管理したりすることがますます重要になっており、そのためには、どのような結果をめざしているかについてネットワークの構成者が共通の認識をもっていなければならない。」と述べ、この枠組みの「目標の設定」「価値観の整合」の重要性を指摘する。

そして「信頼感の醸成」については、O'Leary

and Bingham（2009, p260）が、Van Slyke（2009）の分析結果を踏まえて、ネットワークにおける当事者間の信頼が増すにつれて連携（collaboration）も増し、連携が増すにつれて当事者間の紛争は、対話及び、より非公式な紛争解決手段を通じて解決されるようになる、と指摘する。

このような枠組みに加えて、近年の米国の事例研究（Fedorowicz and Sawyer 2012、Nambisan and Nambisan 2013、Yodsampa 2013、など）の成果として、ネットワーク化が一層深化する中で参加主体間のシナジー効果を生み出すための様々な提案がなされている。例えば、Yodsampa（2013, p32）はアフガニスタンに駐留した米国部隊にて見られた省庁間連携（Interagency Coordination）の実態を分析して、このような提携が成果を上げるために以下のような提案をしている。総括すると、①目標の共有、場の共有、機会の共有、情報の共有、共同での作業といった「共有・協働」の実現、②現場への権限委譲・インセンティブの付与と説明責任のセット、③ファシリテーターや信頼関係の重視など人間関係への配慮、が重視されていると言えよう。

○ 複数機関からなるチームメンバーで場所を共有すると共にメンバー会議を招集すること
○ 定期的に「情報共有」「共同での情報分析・計画立案」の機会をもつこと
○（チームとして）経験あるファシリテーターを任命すること
○ 現場に意思決定権限を委譲すること、協働（coordinate）するインセンティブを設定すること、成果に対する説明責任を設定すること
○ 共有する目標と優先順位とを明確にすること
○ 複数機関から成るチームを組成する際、メンバーの個性やチーム特性に配慮すること

（出所）Yodsampa（2013）

6. ネットワークのガバナンスと業績評価

ネットワークでは複層的な業績評価が求められる一方で、ネットワークであるが故に、ガバナンスに評価が果たす役割が一層重要になるとの側面もある。Moore（2009, p222）は、評価の重要性は単独政府機関の時よりもネットワーク環境にてより一層高まることを指摘した上で、ネットワークでの確固たる評価の枠組みは、単一組織には存在する「中核的な権威機能」（central authority）を代替しうること、そして評価によりネットワークの成果を確認できることは参加者のネットワークへの参画への士気や動機の維持に不可欠なこと、従ってネットワーク型の公共サービスにとり評価制度の運用が極めて重要であることを指摘する。またRabe（2009, p55）は、信頼しうるネットワーク構築のためには、重要な情報が正確にかつ容易にアクセスしうるように参加主体間に流される必要があると指摘する。Koliba（2014, p90）もまた、Moynihan（2008）などを引用しつつ、業績情報を活用したネットワークのガバナンス強化の重要性を指摘する。

またKenis は、ネットワークを維持させて成果の創出を継続させるためには、ネットワークに何らかの「統合機能（integration mechanism）」が備わっていることが必要だとしており、業績情報もその機能を果たすことが可能であると述べている。この「統合機能」とは、ネットワークに参加する多様な利害・関心・立場を持つ多様な主体を結び付けて、ネットワーク全体を機能させる役割、といった意味に理解すれば良いと思われる。[4]

結びにかえて（今後の研究課題）

以上、ネットワークによる公共サービスの提供に関して、そのマネジメント、ガバナンス、業績評価の在り方についての先行研究をレビューしてきた。

ネットワークの形成により、非定型業務や多様性のある業務に柔軟に対応できるようになるなどのメリットがある一方で、多様な利害や関心を有する多様な主体から構成されるネットワークについて、それを円滑にマネジメントすることや、適切なガバナンスを確保することには、単一組織の場合よりも、より一層の複雑さや困難が伴いがちである。例えば、複眼的に森と木とを同時にみる、複数のシステムを同時にまわす、複層的に取

組む、など複数のことに同時に取組むことが求められる。また共有・協働と委譲・裁量など、方向性の異なるプロセスにも同時に取組む必要がある。そして、構成員相互の信頼関係や利害調整の重要性が強調されるのも特徴的である。

このような欧米の先行研究の動向の把握を踏まえ、最後に、本テーマにおける今後の研究課題を整理する。総論としては、例えば以下のような研究課題が抽出できるものと考えられる。

・ネットワークによる公共サービスのマネジメント、ガバナンス、及び業績評価の在り方についての日本の文脈からの研究、すなわち欧米の先行研究の結果との異同についての理論・実証両面からの研究
・成果をあげうるネットワークの構造及びプロセスの研究
・「官官（政府組織間）ネットワーク」と、「官民ネットワーク」におけるマネジメント、ガバナンスや業績評価の在り方の異同の研究
・ネットワークのマネジメント、ガバナンスにおいて「業績情報」が果たす役割の研究
・ネットワークのライフステージの違いに応じたマネジメント、ガバナンスの在り方及びそれに応じた業績評価の在り方の研究
・ネットワークにおける参加者間の利害調整、及びコミュニケーションの在り方の研究
・ネットワークの有効性に対する評価実務

これらの多様な研究課題への取組みは、ネットワークによる公共サービスの重要性が増していく中で、様々な実務的諸課題に対する処方箋の用意につながるものと思われる。そして初期のPFI病院の破綻など、わが国でもネットワークにおける「PPPの失敗」が既に起きていることを踏まえると、これら研究の蓄積は、ひいては「PPPの失敗」の克服への貢献につながるものと考える。筆者自らの努力も含め、今後、様々な研究が蓄積されて研究が前進していくことに期待しつつ、本稿を終えることとしたい。

謝辞

拙稿に対して、複数の査読者の方々から、懇切丁寧なご指摘及び修正のアドバイスを頂戴した。心からの謝意を表したい。それらは本稿の修正プロセスにて、論文の趣旨や意義を明確にしていくための極めて的確かつ強力なガイドとなった。

注記

1 ここで「複数の」とは、複数（2以上）の民間主体により1つの特別目的会社（SPC）を組成するケースを含め、合計3以上の主体の参画があるケースを想定している。
2 PFIの場合、個々の事例の状況（独立採算型かサービス購入型か、BTOかBOTか、などを含む）により、本稿でいうネットワーク性には高低がある。また民間委託や指定管理者制度の事例の中にも、民が事業リスクを負う場合など、本稿でいうネットワークに該当するものがあると想定される。
3 Popp et al.（2014, p79-80）も指摘するように、ネットワークにおける成果の発生プロセスは、単純なロジックモデル（投入⇒結果⇒成果）のようなリニアなものではないと考えられるが、それでも、ロジックモデルの考え方はネットワークの業績評価においても非常に有用であると考えられる。「if-then」の関係でネットワーク全体を整理することで、多数の関係者によるネットワークの業務と成果発生の間のメカニズムの全体像が整理できるようになる。
4 Public Management Research Conference 2015（Public Management Research Association）におけるPatrik Kenis, Bart Cambre, and Steven van den Oord による報告「What Holds Inter-organizational Networks Together? A Study of Effective Integration Mechanism of Inter-organizational Networks」、及び当該Conferenceにおける筆者との議論による。

参考文献

Agranoff, Robert（2007）. *Managing within Networks: Adding Value to Public Organizations.* Washington, D.C.: Georgetown University Press.

Agranoff, R.（2008）. Enhancing Performance Through Public Sector Networks: Mobilizing Human Capital in Communities of Practice. *Public Performance & Management Review.* Vol. 31, No. 3, March 2008. 320-

347.

Agranoff, R.（2012）. *Collaborating to Manage: A Primer for the Public Sector.* Washington, D.C.: Georgetown University Press.

Agranoff, Robert（2014）. Bridging the Theoretical Gap and Uncovering the Missing Holes. In Robyn Keast, Myrna Mandell, and Robert Agranoff（eds.）, *Network Theory in the Public Sector: Building New Theoretical Frameworks.* New York, New York: Routledge. 193-209.

Cepiku, Denita（2014）. Network Performance: Toward a Dynamic Multidimensional Model. In Robyn Keast, Myrna Mandell, and Robert Agranoff（eds.）, *Network Theory in the Public Sector: Building New Theoretical Frameworks.* New York, New York: Routledge. 174-189.

Donahue, John D., and Zeckhauser, Richard J.（2011）. *Collaborative Governance.* Princeton, N.J.: Princeton University Press.

Eggers, William D.（2009）. From Conflict to Collaboration: Lessons in Networked Governance from the Federal Cooperative Initiative. In Stephen Goldsmith and Donald F. Kettl（Eds.）, *Unlocking the Power of Networks: Keys to High-performance Government.* Washington, D.C.: The Brookings Institution, 15-33.

Fedorowicz, Jane, and Sawyer, Steve（2012）. *Designing Collaborative Networks: Lessons Learned from Public Safety.* Washington D.C.: IBM Center for The Business of Government.

Frederickson, D., and H. G. Frederickson（2006）. *Measuring the Performance of the Hollow State.* Washington, D.C.: Georgetown University Press.

Goldsmith, Stephen, and Eggers, William D.（2004）. *Governing by Network: The New Shape of the Public Sector.* Washington, D.C.: The Brookings Institution.（スティーブン・ゴールドスミス、ウィリアム・D・エッガース著、城山英明・奥村裕一・髙木聡一郎監訳『ネットワークによるガバナンス－公共セクターの新しいかたち』学陽書房．2006年）

Hale, K.（2011）. *How Information Matters: Networks and Public Policy Innovation.* Washington, D.C.: Georgetown University Press.

Isett, Kimberley R., Mergel, Ines A., LeRoux, Kelly, Mischen, Pamela A., and Rethemeyer, R. Karl（2011）. Networks in Public Administration Scholarship: Understanding Where We Are and Where We Need to Go. *Journal of Public Administration Research and Theory.* Vol.21: i157-i173.

Kenis, Patrick P., and Provan, K. G.（2009）. Towards an Exogenous Theory of Public Network Performance. *Public Administration.* Vol. 87, No.3, 2009. 440-456.

Kettl, Donald F.（2009a）. *The Next Government of the United States: Why Our Institutions Fail Us and How to Fix Them.* New York, N.Y.: W.W. Norton & Company, Inc..（ドナルド・ケトル著、稲継裕昭監訳、浅尾久美子訳『なぜ政府は動けないのか－アメリカの失敗と次世代型政府の構想』勁草書房．2011年）

Kettl, Donald F.（2009b）. The Key to Networked Government. In Stephen Goldsmith and Donald F. Kettl（Eds.）, *Unlocking the Power of Networks: Keys to High-performance Government.* Washington, D.C.: The Brookings Institution, 1-14.

Klijn, Erik-Hans（2010）. Trust in Governance Networks: Looking for Conditions for Innovative Solutions and Outcomes. In Stephen P. Osborne（Ed.）, *The New Public Governance?: Emerging Perspectives on the Theory and Practice of Public Governance.* Oxon, United Kingdom: Routedge, 303-321.

Koliba, C., Campbell, E., and Zia, A.（2011）. Performance Measurement Consideration in Congestion Management Networks: Evidence From Four Cases. *Public Performance Management Review.* 34（4）: 520-548.

Koliba, Christopher, Meek, Jack W., and Zia, Asim（2011）. *Governance Networks in Public Administration and Public Policy.* Boca Raton, FL.: CRC Press.

Koliba, Christopher（2014）. Governance Network Performance: A Complex Adaptive Systems Approach. In Robyn Keast, Myrna Mandell, and Robert Agranoff（eds.）, *Network Theory in the Public Sector: Building New Theoretical Frameworks.* New York, New York: Routledge. 84-102.

Koppenjan, J.（2008）. Creating a Playing Field for Assessing the Effectiveness of Network Collaboration by Performance Measures. *Public Management Review.* Vol. 10, Issue 6, 2008. 699-714.

Mandell, M. P., and Keast, R.（2008a）. Introduction. *Public Management Review.* Vol. 10, Issue 6, 2008. 687-698.

Mandell, M. P., and Keast, R.（2008b）. Evaluating the

Effectiveness of Interorganizational Relations through Networks: Developing a Framework for Revised Performance Measures. *Public Management Review.* Vol. 10, Issue 6, 2008. 715-731.

McGuire, Michael, and Agranoff, Robert（2014）. Network Management Behaviors: Closing the Theoretical Gaps. In Robyn Keast, Myrna Mandell, and Robert Agranoff（eds.）, *Network Theory in the Public Sector: Building New Theoretical Frameworks.* New York, New York: Routledge. 137-156.

Meier, K.J., and O'Toole, L. J.（2003）. Public Management and Educational Performance: The Impact of Managerial Networking. *Public Administration Review.* 63（6）: 689-99.

Milward, H. Brinton, and Provan, Keith G.（2006）. *A Manager's Guide to Choosing and Using Collaborative Networks.* Washington D.C.: IBM Center for The Business of Government.

Moore, Mark H.（2009）. Networked Government: Survey of Relations, Forms, and Techniques. In Stephen Goldsmith and Donald F. Kettl（Eds.）, *Unlocking the Power of Networks: Keys to High-performance Government.* Washington, D.C.: The Brookings Institution, 190-227.

Moynihan, D. P.（2008）. *The Dynamics of Performance Management: Constructing Information and Reform.* Washington, D.C.: Georgetown University Press.

Nambisan, Satish, and Nambisan, Priya（2013）. *Engaging Citizens in Co-Creation in Public Services: Lessons Learned and Best Practices.* Washington, D.C.: IBM Center for The Business of Government.

O'Leary, Rosemary R., and Bingham, L. B.（2009）. Surprising Findings, Paradoxes, and Thoughts on the Future of Collaborative Public Management Research. In Rosemary O' Leary, and Lisa Blomgren Bingham（Eds.）（2009）. *The Collaborative Public Manager: New Ideas for the Twenty-First Century.* Washington, D.C.: Georgetown University Press. 255-269.

Osborne, Stephen P.（2010a）. Introduction: The（New）Public Governance: A Suitable Case for Treatment?. In Stephen P. Osborne（Ed.）, *The New Public Governance?: Emerging Perspectives on the Theory and Practice of Public Governance.* Oxon, United Kingdom: Routedge, 1-16.

Osborne, Stephen P.（2010b）. Conclusions: Public Governance and Public Service Delivery: A Research Agenda for the Future. In Stephen P. Osborne（Ed.）, *The New Public Governance?: Emerging Perspectives on the Theory and Practice of Public Governance.* Oxon, United Kingdom: Routedge, 413-428.

Poister, Theodore H.（2010）. Performance Measurement: Monitoring Program Outcomes. In Joseph S. Wholey, Harry Hatry, and Kathryn E. Newcomer（Eds.）, *Handbook of Practical Program Evaluation,* Third Edition. San Francisco, California, United States: Jossey-Bass, 100-124.

Popp, Janice K., Milward, H. Brinton, MacKean, Gail, Casebeer, Ann, and Lindstrom, Ronald（2014）. *Inter-Organizational Networks: A Review of the Literature to Inform Practice.* Washington, D.C.: IBM Center for The Business of Government.

Provan, Keith G., and Kenis, Patrik（2007）. Modes of Network Governance: Structure, Management, and Effectiveness. *Journal of Public Administration Research and Theory.* 18:229-252

Provan, Keith G., and Lemaire, Robin H.（2012）. Core Concepts and Key Ideas for Understanding Public Sector Organizational Networks: Using Research to Inform Scholarship and Practice. *Public Administration Review.* September/October 2012, Vol. 72, No.5:638-648.

Provan, Keith G., and Milward, H. Brinton（1995）. A Preliminary Theory of Interorganizational Network Effectiveness: A Comparative Study of Four Community Mental Health Systems. Administrative Science Quarterly 40（1）: 1-33.

Provan, Keith G., and Milward, H. Brinton（2001）. Do Networks Really Work? A Framework for Evaluating Public-Sector Organizational Networks. *Public Administration Review.* July/August 2001, Vol. 61, No.4: 414-423.

Rabe, Barry E.（2009）. Governing the Climate from Sacramento. In Stephen Goldsmith and Donald F. Kettl（Eds.）, *Unlocking the Power of Networks: Keys to High-performance Government.* Washington, D.C.: The Brookings Institution, 34-61.

Radin, B.（2006）. *Challenging the Performance Movement:*

Accountability, Complexity, and Democratic Values. Washington, D.C.: Georgetown University Press.

Turrini A., Cristofoli D., Frosini, F., and Nasi G.（2010）. Networking Literature about Determinants of Network Effectiveness. *Public Administration.* Vol. 88, No.2, 2010. 528-550.

Van Sleke, David M.（2009）. Collaboration and Relational Contracting. In Rosemary O'Leary, and Lisa Blomgren Bingham（Eds.）（2009）. *The Collaborative Public Manager: New Ideas for the Twenty-First Century.* Washington, D.C.: Georgetown University Press. 137-156.

Voets, J., Van Dooren, W. and De Rynck, F.（2008）. A Framework for Assessing the Performance of Policy Networks. *Public Management Review.* Vol.10, Issue 6, 2008. 773-790.

Voets, Joris (2014). Developing Network Management Theory through Management Channels and Roles. In Robyn Keast, Myrna Mandell, and Robert Agranoff (eds.), *Network Theory in the Public Sector: Building New Theoretical Frameworks*. New York, New York: Routledge. 118-134.

Yodsampa, Andrea Strimling（2013）. *Coordinating for Results: Lessons from a Case Study of Interagency Coordination in Afghanistan.* Washington D.C.: IBM Center for The Business of Government.

（2016.5.2 受理）

Networked Public Services: Its Management, Governance, and Performance Measurement -Reviews of the Literature and Agenda for the Future-

Jun Arakawa
Doctoral Program, Graduate School of Media and Governance, Keio University
arakawa525jun@gmail.com

Masatoshi Tamamura
Professor, Faculty of Policy Management, Keio University

Abstract

This paper discusses management and governance system of networked public services, and its performance measurement and management system. In particular, the paper aims to review the literature on this issue mainly in western countries in order to overview the discussion.

There has been cases where networks are formulated between government agencies or between government and private organizations such as corporation or non-profit organization, in order to better respond to complicated policy issues and/or various policy needs. By formulating networks, certain merits could be obtained such as more flexible reactions for non-routine, dynamic and diverse policy agenda. At the same time, however, more complicated and often difficult management and governance are needed than public service provided by a single organization. For example, multiple system, and multi-layer approaches are becoming important for network management and governance. In addition, different approaches should be simultaneously introduced in the networks such as corporation/collaboration/sharing and delegation/discretion. Furthermore, mutual trust among network participants, and coordination of various interests, including conflict management are also important for better network management and governance.

For the future, more research, both theoretical and empirical ones, are expected to be conducted in Japan on this matter, and they could contribute to practical solution of problems related to network performance as well as discussion under global environment, and then to overcome “PPP failures”.

Keywords

network, management, governance, performance measurement, performance information

日本評価学会春季第 12 回全国大会
「新開発協力大綱下での評価」
開催の報告とお礼

2015年5月30日開催（於：JICA地球ひろば）の日本評価学会春季第12回全国大会には98名の方々にご出席いただきました。誠にありがとうございました。会員各位の日頃の研究や実践活動の報告を基に、評価研究者、実務者の間の経験、情報、知識を共有化する場として、皆様にとって有益な機会となりましたら幸甚に存じます。多くの方のご尽力により本大会を開催することができましたこと、心より感謝申し上げます。今後とも当学会の活動にご高配賜りますようどうぞ宜しくお願い申し上げます。

実行委員長　　和田義郎（国際協力機構）
プログラム委員長　牟田博光（東京工業大学）

2015年5月30日（土）

10：00 - 10：30	受　付		
午前の部 10：30 - 12：30	**共通論題1** 評価における客観性 ―課題と展望― （山谷清志） Rm.600	**自由論題1** 参加型評価 （林薫） Rm.201AB	**自由論題2** 評価の枠組み （佐々木亮） Rm.202AB
お昼休み 12：30 - 13：45	お昼休み （12：35 - 13：30　理事会）		
午後の部Ⅰ 13：45 - 15：45	**共通論題2** 新しい開発協力大綱での評価 （和田義郎） Rm.600	**共通論題4** 評価に関わる関係者のキャパシティ・ビルディングとエンパワーメント （源由理子） Rm.201AB	**自由論題3** 政策評価 （小野達也） Rm.202AB
午後の部Ⅱ 16：00 - 18：00	**共通論題3** 評価のガバナンス （廣野良吉） Rm.600	**共通論題5** 評価における科学性： エビデンスの実践的活用とその方向性 （正木朋也） Rm.201AB	

（　　）は座長です。

日本評価学会春季第 12 回全国大会
「新開発協力大綱下での評価」プログラム詳細

5 月 30 日（土）受付 10:00 - 10:30

5 月 30 日（土）午前の部 10:30 - 12:30

共通論題1　評価における客観性　―課題と展望―　Rm.600

座長　山谷　清志　同志社大学

K1-1	総務省の客観性担保評価活動	○	宗高　有吾	総務省
K1-2	独立行政法人評価の客観性担保		森田　弥生	筑波大学大学院
K1-3	評価と監査の理論上の区別		橋本　圭多	同志社大学
K1-4	政府の評価における客観性の意味とその変遷		山谷　清志	同志社大学

自由論題1　参加型評価　Rm.201AB

座長　林　薫　文教大学

J1-1	学生経験調査を利用した大学生の能力に関わる分析 ―企業が重視する能力と大学生の主観的な能力評価の比較―	○	和嶋　雄一郎	大阪大学
			中村　将大	大阪大学
			齊藤　貴浩	大阪大学
J1-2	効果的プログラムモデル形成のための実践家参画型評価アプローチ法の開発（その7）；効果モデル改善ステップにおける「全国試行評価調査（1年間の提示版プログラム試行調査）」の位置とその検証	○	新藤　健太	日本社会事業大学
			大島　巌	日本社会事業大学
			植村　英晴	日本社会事業大学
			浦野　由佳	日本社会事業大学
			方　真雅	日本社会事業大学
			全　形文	日本社会事業大学
			村里　優	日本社会事業大学
			石井　真紀子	日本社会事業大学
			益子　徹	日本社会事業大学
			巣立　佳宏	埼玉工業大学
J1-3	フランスの学校外部評価 ―目標管理型評価と参加型評価、2つの評価手法の意義と可能性についての考察―		佐々木　織恵	東京大学（博士課程）

自由論題2　評価の枠組み　Rm.202AB

座長　佐々木　亮　国際開発センター

J2-1	イギリスにおける研究力評価の動向―REF2014のImpact評価を中心に―		藤井　翔太	大阪大学
J2-2	環境配慮行動誘発に向けた効果的アプローチの検証 モザンビークにおけるランダム化比較試験の事例	○	青柳　恵太郎	国際協力機構／東京大学
			細野　智之	日本工営株式会社
J2-3	セオリー評価における「理論」とは何か		永井　隆夫	同志社大学大学院（博士課程）
J2-4	学校評価計画のための論理枠の系統的作成方法についての一考察 ～ ‘Rainbow Framework’ の紹介とその活用方法～		橋本　昭彦	国立教育政策研究所

5 月 30 日（土）お昼休み 12:30-13:45（12:35 - 13:30　理事会）

5 月 30 日（土）午後の部Ⅰ 13:45 - 15:45

共通論題2　新しい開発協力大綱での評価　Rm.600

座長　和田　義郎　国際協力機構
コメンテーター　清原　剛　外務省
廣野　良吉　成蹊大学
牟田　博光　国際開発センター
鴨谷　哲　国際協力機構

K2-1	開発協力大綱における評価		清原　剛	外務省
K2-2	簡単な論点整理	○	和田　義郎	国際協力機構

共通論題4　評価に関わる関係者のキャパシティ・ビルディングとエンパワーメント　Rm.201AB

座長　源　由理子　明治大学

K4-1	評価のキャパシティ・ビルディングとエンパワーメント		山谷　清志	同志社大学
K4-2	学校評価における教員のキャパシティ･ビルディング過程と教育改善効果の検証:学校評価実践支援の実践事例から		池田　琴恵	東京福祉大学
K4-3	福祉サービス利用における当事者のエンパワメント		藤島　薫	東京福祉大学

自由論題3　政策評価　Rm.202AB

座長　小野　達也　鳥取大学

J3-1	独立財政機関（IFIs）の活動内容とその評価：改訂版		上野　宏	統計研究会／国際開発センター
J3-2	知事による政策の自己評価の分析		本田　正美	東京大学大学院
J3-3	自治体行政評価の改善研究		高千穂　安長	ノースアジア大学
J3-4	オバマ政権の社会政策における挑戦とCBO 40年の意味するもの		上野　真城子	アジア都市コミュニティー研究センター（UCRCA）

5 月 30 日（土）午後の部Ⅱ 16:00 - 18:00

共通論題3　評価のガバナンス　Rm.600

座長　廣野　良吉　成蹊大学

K3-1	自治体評価のガバナンス		田中　啓	静岡文化芸術大学
K3-2	日本の政策分析･評価ガバナンスの評価と独立財政機関（IFIs）		上野　宏	統計研究会／国際開発センター
K3-3	ネパールにおける評価ガバナンスについて：課題と展望		佐々木　亮	国際開発センター

共通論題5　評価における科学性：エビデンスの実践的活用とその方向性　Rm.201AB

座長　正木　朋也　東京大学大学院

K5-1	保健医療領域における評価の動向：研究の多様化と指針整備	○	正木　朋也	東京大学大学院
K5-2	国際開発評価における透明性確保の取り組み		青柳　恵太郎	東京大学大学院
K5-3	教育領域における政策・実践でのエビデンス活用の課題		岩崎　久美子	国立教育政策研究所
K5-4	社会分野における昨今の「エビデンス」活用に関する批判的考察 SROIをめぐって		津富　宏	静岡県立大学

共通論題セッション報告

共通論題セッション１「評価における客観性ー課題と展望ー」

座長　山谷 清志（同志社大学）

政策評価、独立行政法人評価、あるいはODA評価、または男女共同参画政策の評価など、さまざまな評価がどのようにその客観性を担保すべきか、担当者は頭を痛めてきた。ひとつの方法としては評価の手続き的側面に注目し、外部の第三者による評価や政策実施機関から離れた議会･国会が行う評価のように制度として対応する場合がある。他方、科学的な知見や手法を用いて、実際のエビデンスにもとづく評価、あるいはプロの専門家による専門評価として客観性を確保しようとする場合もある。ただし、この場合、社会科学の応用の意味を知っているエキスパート（学位や専門資格を持ち実務経験・教育経験がある人）が必要になる。そしてそれは実務では難しい。

そこで手間をかけたチェックを繰り返すことになる。日本ではここから3つ問題が発生した。①複数実施する評価の重層化が複雑化と混乱を招き、②評価担当者（素人）を消耗させながら役に立たない評価を垂れ流し、③評価を知らない人（トップや議員）が監査の意味合いで詳細な評価を求め「重箱の隅をほじくる」である。

しかし、そもそも、日本では客観的に評価を進めるとはどういったことなのか、そして評価それ自体の客観性とはなにか、実はよく知られていない。そこで政策評価と独立行政法人評価の2つから、実務における評価の歴史と現状、そして評価の困難さを紹介し、その後に理論的な展望を探るため、この共通論題は設定された。

第1報告者の宗高有吾氏（総務省行政評価局）からは、総務省行政評価局による客観性担保評価活動について報告があった。客観性担保評価とは各府省が実施した政策評価を総務省行政評価局が点検する活動を指し、これは「行政機関が行う政策の評価に関する法律」（2001年）によって開始されたため10年以上の歴史がある。報告ではこれまでの点検活動の取組を俯瞰し、各府省の評価の実施状況を網羅的に確認する「審査」や、より評価の内容に踏み込んで確認する「認定関連活動」を紹介した。客観性担保評価活動における評価の客観性・厳格性の基本的な捉え方は、その実施が開始された2002年から基本的に一貫しているが、これまでの取組における点検の着眼点を整理・比較すれば、各評価対象分野において求められる客観性・厳格性はそれぞれ微妙に異なることが分かる。

第2報告者の森田弥生氏（筑波大学大学院、前・総務省行政評価局）からは独立行政法人制度とその評価、そして3度目の国会提出でようやく成立した独立行政法人通則法改正法（2015年4月スタート）について報告があった。今般の改正内容を通じて、どのように独立行政法人評価の客観性が担保されているのか、それが論点である。そもそも独立行政法人は、公費を投入されるため国民へのアカウンタビリティが問われる存在であるが、時代背景などにより問われるアカウンタビリティが異なる。そこでこの報告では、①パフォーマンス・アカウンタビリティ、②コンプライアンス・アカウンタビリティの2つに注目し、これらの視点から、独立行政法人評価における客観性担保について分析をした。制度内容の一つであるPDCAが機能する目標・評価の仕組みの再構築は、パフォーマンス・アカウンタビリティの側面が多い一方で、独立行政法人の内外から業務運営を改善する仕組みの導入は、コンプライアンス・アカウンタビリティの側面が多いと考えられる。これらを包含する評価の客観性を担保するツールとして期待されるのは、①評価スキーム、②評価基準、③目標設定である。新制度が3つをどのように実行あるものにできるのか、それが課題である。

第3は橋本圭多氏（同志社大学）の評価と監査の理論上の区別に関する報告である。政府部門の評価に

おいて客観性が議論される原因として、評価における監査様式の混入が指摘される。本来、評価と監査は理論上区別される活動であるが、評価実務では両者の混同から区別が曖昧になってきたのである。その背景には、メタ評価における基準の存在や監査機能の発展がある。本来、外部監査は最高会計検査機関が担うのが国際的には一般的であるが、日本ではこれら機能の一部を総務省行政評価局が担っていると報告者は考える。それは、総務省行政評価局が行う客観性担保評価活動や目標管理型政策評価、独立行政法人評価、行政評価局調査（統一性・総合性確保評価と行政評価・監視）が先の監査類型に適合しているからである。また、そもそも、外部監査の目的はアカウンタビリティの確保であるが、日本の政策評価も「行政機関が行う政策の評価に関する法律」が第1条で「国民に説明する責務」を規定しているが、この点も日本で評価と監査が混同される原因の一つになっている。ただし、評価においてアカウンタビリティが強調されるようになったのはアメリカの最高会計検査機関（GAO）がプログラム評価を採用したからであり、評価におけるアカウンタビリティ確保が可能となるのもアメリカの厳格な三権分立に基づく統治制度に依拠している背景を無視していることが日本では多い。したがって、もし総務省行政評価局がアカウンタビリティ確保評価の実効性を高めるのであれば、評価対象から独立した知識基盤を獲得し「情報源を依存する状態」を回避し、アカウンタビリティが確保されない場合の制裁発動などの課題を解決しなければならない。それができない限り、日本の政策評価は情報公開としての役割にとどまらざるを得ないと言う。

第4報告、山谷清志（同志社大学）は、以上の3報告をふまえ評価における客観性の議論が日本ではひとつの特徴を持っていると指摘する。すなわち、評価対象を所管する原課が評価そのものも担当するために発生する「お手盛り評価」をどのように避けるべきか、そのため客観性が話題に上ったという特徴である。公共事業の採択や不採択、各種基本計画の妥当性チェック、あるいは福祉事業の有効性評価まで、日本では客観性の議論がいまだに多いのは、このためである。他方、諸外国では1980年代には「評価のポリティックス」の文脈で議論されたことがあるが、21世紀になってからは評価の客観性･主観性の議論は少ない。そこで再度評価の客観性とは何か、またそれはなぜ、どのような文脈で問題視され、その問題をどうやって解決するのか（あるいはできないのか）について議論したのが、第4報告である。

評価をめぐる客観性の議論は、日本の政治と行政の制度と文化に大きく影響を受け、諸外国の評価の実践とはその本質において異なる使い方をされてきている。その典型は政策評価や独立行政法人の評価と監査と、行政事業レビューとの関係である。一部の研究者がその理論と実践について指摘するだけであるが、その実態は研究者の間でも国民や報道機関の間でもあまりよく知られていない。この実態や実践においても、客観性が大きく問題になるはずであるが、それは今後の日本評価学会で議論されるべきテーマになるであろう。

共通論題セッション２「新しい開発協力大綱での評価」

座長　和田 義郎（国際協力機構）
討論者　清原 剛（外務省）
廣野 良吉（成蹊大学）
牟田 博光（国際開発センター）
鳴谷 哲（国際協力機構）

ODA評価において、数多く議論されているのは、ODA事業の評価であるが、本共通論題では、昨年度に制定された新しい開発協力大綱について、ODA政策の評価という観点での議論が行われた。開発協力大綱における新しい論点や運用、そのインパクトとともに、新しい開発協力大綱が、これまでのODA大綱の評価をどのように反映しているのか、という論点も提出された。セッションは、最初に報告者より、新大綱についての説明がなされ、その後、論点整理を経て、討論者によるコメント、質疑応答、フロアとの意

見交換という形で進められた。

報告者の清原剛氏（外務省）「開発協力大綱における評価」においては、大綱改定の背景、改定プロセス、大綱の特徴、評価へのインプリケーションなどについての発表があり、特に、大綱の特徴として、戦略性、ODAの位置づけの変化などの論点が述べられた。

第1討論者の廣野良吉氏（成蹊大学）よりは、大綱の戦略性についてのコメントが述べられた。

第2討論者の牟田博光氏（国際開発センター）よりは、「ODA大綱から開発協力大綱へ」として、過去のODA評価案件のレビューによるODA大綱の評価についての発表があった。2003年から2013年度に実施した外務省ODA評価67件を対象に、ODA大綱を切り口としたメタ評価を行い、提言・教訓515件の分類・体系化を行った結果、ODA大綱の基本方針との整合性、プロセスの適切性、戦略性の明確化がとくに数多く提言された点などが確認された。これらの点が、新開発協力大綱に反映され、政策→評価→次の政策決定への活用、というサイクルが実践されている点が重要としている。

第3討論者の鴫谷哲氏（国際協力機構）「開発協力大綱とJICAの評価業務について」においては、開発協力大綱の理念である国益の確保、重点施策としての質の高い成長、フィードバック、事情の考慮、外交的視点などの実施方策に対して、どのようにJICAの評価が対応してゆくか、その方向性について議論された。

その後、新大綱の決定プロセスについての質疑応答、新大綱の適用についての議論などがフロアからも提出され、活発な議論が行われた。ODAの政策は、外交政策や対外経済政策、国内の経済事情など、数多くの要素を反映して決定され、実施されるものであり、極めて複雑な構造を持つものであるが、開発協力大綱という形で政策が発表され、国民の間で議論、共有されることで、その理解が進むことは、極めて重要である。ODA政策の評価は、ODAを改善してゆく上で必要不可欠であることが、今回の共通論題で改めて認識されたと考える。

共通論題セッション３「評価のガバナンス」

座長　廣野 良吉（成蹊大学）

今日では、一方でグローバル化の進展に伴い、評価対象があらゆる国、あらゆる分野、あらゆる地域（国内）に及んでいるだけでなく、これらの国・分野・地域にまたがって複合化しており、多様な専門領域・分野をもった専門家チームによる評価体制の導入を通じて、一層の科学性・客観性を担保する仕組みが急務となっている。他方では、市民意識・社会の急速な台頭と納税者意識の向上により、さらに財政赤字の大幅な増大に伴う無駄な政府支出に対する国民、市民、住民の監視の目が強化されてきた。かくして、中央政府、地方自治体を問わず、行政機関の政策、施策、事務事業等の客観的評価による効率性、有効性と共に、透明性と負託責任（accountability）への各界の要求が顕著となってきている。こうして、我が国でも「管理手段」としての評価活動は、評価手法・技術の改善、データの精緻化等によって評価の質を高めて、評価対象の政策等の改善を図りつつある。しかし、評価のガバナンスの未整備から起因する「評価の恣意性と客観性の欠如」という、もう一つの評価の課題が途上国のみならず、我が国を含めた先進諸国でも存在する。わが国で2001年に制定された「政策評価法」は、政策、プログラム等の評価とその公表を義務付けたが、それは「管理手段」としての評価の徹底以外に、正に行政機関の透明性、負託責任を明確にするための「評価のガバナンス体制」の改善のためでもあった。しかし、これらの制度設計が現実問題として、「評価のガバナンス」の改善に有効に機能しているのかどうか、評価の客観性の確保と評価に対する国民、市民、住民の信頼性の改善に寄与しているのかどうかが、今真剣に問われている。本セッションでは、3人の評価研究者がそれぞれ、我が国における国レベル、地方自治体レベルの評価のガバナンスと、ネパールにおける評価のガバナンスについて発表し、参加者と率直な質疑討論がなされた。

本セッションの最初の報告者である田中啓氏（静岡文化芸術大学）は、日本の「自治体評価のガバナンス」と題する発表で、①ガバナンスの概念を「権限を持った主体 principal, 主権者）と主権者の委託により行動を行う主体（agent, 執行者）が別々に存在し、主権者が執行者の行動を統制すること」という定義（赤井伸郎、2006）を援用して、②自治体が評価を適切に実施するような統制の在り方として「市民統治モデル」と「首長統治モデル」という2つの「類型」を示し、それぞれのガバナンスモデルの特性を論じ、③我が国の地方自治体の評価ガバナンスモデルの大半は「首長統治モデル」であり、④H市における自治体評価のガバナンスの実例を取り上げて、当該市だけでなく、広く一般的に「市民統治モデル」で想定されたガバナンス・システムは監視役不在でほとんど機能せず、「首長統治モデル」も形式上は体制が整っているものの、現実には首長の関心度の違いや職員の技術的困難から機能せず、一方で透明性・コミットメントの欠如、他方で評価の手抜きや粉飾の横行が現実の問題点であるという、鋭くかつ厳しい批判がなされた。しかし、質疑応答でも指摘されたように、⑤主権者である住民、その代表である首長や市議会や専門家による第三者評価機関の関与・監視を高めることにより、自治体の評価ガバナンス・システムの改善が重要であるという結論になった。

報告の第2は、上野宏氏（統計研究会／国際開発センター）による「日本の政策分析・評価ガバナンスの評価と独立財政機関（IFIs）であり、①ガバナンスの定義が不十分という認識に基づき、ガバナンス評価のための一般的規範基準19項目を設定し、②政策分析・評価活動のガバナンスを査定する査定基準（評価へのコミットメントや独立性等）を作成し、③この査定基準委に基づき我が国の政策分析・評価の査定を行うと「日本の評価制度は非常に多くの問題を抱えている」と指摘し、特に、政策の改善、国民市民への覚醒、行政改革、財政健全化等への貢献は皆無に等しいと結論付けた。さらに、④この査定の結果として、国会による独立財政機関の設立の必要性を力強く説いたている。この課題については、衆参両議院の予算委員会や行政監視委員会による審議等、国民の代表である国会による行政活動の監視・統治が行われてきたが、いずれも機能しない中で、2012年日本評価学会も筆者が委員長となって予算審議改革検討委員会を設置、予算審議改革に関する提言を纏めたが、理事会内の一部の反対で提言の実現は不可能であった。その後さらに一部国会議員が超党派で類似の独立予算分析機関の設立を推奨し、その草案を準備してきたが、昨今の戦後安全保障体制の根本的変革という大きな政治的課題の出現により、本課題へ対処する法案は未だに国会へ上程されていない。

報告の第3は、佐々木亮氏（国際開発センター）による「ネパールにおける評価ガバナンスについて：課題と展望」があった。ネパールでは、他の多くの途上国と同様に、伝統的な半封建的社会構造の下では、「評価の文化」の普及は今後長年の課題であるが、国内外の民主化を求める政治のグローバル化の中で漸くプロジェクト評価への関心が高まり、GTZ、JICAや国際機関等の協力の下で今世紀になって評価体制の構築が始まった。中央政府国家計画委員会（NPC）事務局、各省庁、群との連携の中で、制度設計が進み、「評価の用語とスタンダード」や「国家M&Eガイドライン」の策定、義務付けられたM&Eを通じた計画策定そのものへの反映が試みられているが、M&E研修の定期化、予算的裏付け、ネパール評価学会の活性化等多くの課題を抱えており、適切な評価ガバナンスの構築は長い道のりとなるであろうという指摘も報告後の討論で示された。JICAの支援の下で日本評価学会が、ネパールの評価体制の確立に一役かっていることは、当事者政府はもちろんのこと、アジア太平洋評価協会（APEA）や国際開発評価協会（IDEAS）も高く評価していることを付言したい。

3人の報告後の参加会員相互の論議では、日本評価学会が、今回初めて「評価のガバナンス」を正面から共通論題として取り上げたことを高く評するだけでなく、わが国におけるこの課題への積極的な役割を期待する意見が表明された。

共通論題セッション４「評価に関わる関係者のキャパシティ・ビルディングとエンパワーメント」

座長　源 由理子（明治大学）

多様な価値観が存在する現代社会において、価値づけの行為である評価には、その実態に合わせて多様な方式が存在するはずである。とりわけ、あらかじめ設定された価値基準（目標）に基づく評価専門家による科学的調査だけではなく、ステークホールダーが評価に関わることで、多元的な価値が反映される評価のあり方が問われているのではないか。そのためには、評価に関わる人々のキャパシティも問われる。このような認識のもと、本セッションでは、これまで評価の三大目的としてあげられてきた「アカウンタビリティの確保」、「マネジメントの支援」、「専門分野における知的貢献」とは異なるもうひとつの評価目的である「評価に関わる人々のキャパシティ・ビルディング（能力構築）やエンパワーメント」の今日的意義やその課題について議論することを目的とした。

第一報告者の山谷清志会員からは、「評価のキャパシティ・ビルディングとエンパワーメント」と題し、フェッターマンのエンパワーメント評価を政策リテラシーならびに民主主義との関連から考察したエンパワーメントの内容と中間支援団体の位置づけについて発表があった。評価が民主主義のツールであるとする言説はよく聞かれるが、果たしてそれがどのように実現され得るのかについての問題提起は、今後の日本における評価学のあり方を問う上でも刺激的な問題提起であった。

第二報告書の池田琴恵会員からは、「学校評価における教員のキャパシティ・ビルディング過程と教育改善効果の検証」について、GTO（Getting to Outcomes™）のアプローチを使った学校評価の事例をとおし、評価を手段としたキャパシティ・ビルディングがどのように教育改善に役立つのかの報告があった。特に、キャパシティ・ビルディング（能力構築）とエンパワーメント概念の違いを明らかにしながら、プロセスであるエンパワーメントと結果としてのエンパワーメントの二つの位相から分析した興味深い論考であった。

第三報告書の藤島薫会員からは、「福祉サービス利用における当事者のエンパワーメント」について、福祉分野におけるソーシャルワークの評価に当事者を巻き込んだ参加型評価を取り入れることの意義について報告があった。参加型評価の目的のひとつである当事者の主体性とエンパワーメントが評価対象事業そのものの目的と合致している事例であるが、より効果をあげるために解決指向アプローチを導入した貴重な実践報告であった。

これらの報告に対し、フロアーからは、キャパシティ・ビルディングとエンパワーメントの関係、評価をする際の評価される側の負担感、NPOなどの中間支援評価団体の現状、エンパワーメント評価の対象者の範囲など、エンパワーメントを目的とした評価の実践について質問・意見が提示された。また、このような評価にどこまで「評価」という言葉を使うことが適切なのかについても問題提起があり、利害関係者を巻き込んだ実社会に根付く評価のあり方について意見交換がなされた。第四の評価の目的である「キャパシティ・ビルディングとエンパワーメント」を検討する上で多くのヒントを得ることができたセッションであった。

共通論題セッション５「評価における科学性：エビデンスの実践的活用とその方向性」

座長　正木 朋也（東京大学大学院）

当学会社会実験分科会発足の2002年以来、分科会メンバーを中心に2006および2010年の当学会誌の特集や関係する企画セッションなどを通してEBP（Evidence-based Practice）の普及啓蒙がなされている。社会

実験分科会発足当時、国内では保健医療領域外のRCT（Randomized Controlled Trial）の報告は殆ど見当たらなかった。しかし、その後、関係諸氏の努力により様々なプログラムや施策への応用もみられるようになってきた。

本セッションではこうした状況を鑑み、エビデンス（科学的に根拠のある証拠）の産出・伝達とその実践的活用について、保健医療、国際協力、教育、および社会科学の4分野の現状を共有し、その課題と解決策について「評価の科学性」とは何かということも念頭に議論を行った。

冒頭で先ず座長よりこのセッションが持たれた背景と趣旨説明および論点提起を行い、客観性と科学性についての整理とともに、保健医療分野のその後に著変はないことを概括した。

次いで第1報告者の青柳恵太郎氏（東京大学大学院）からは、国際開発評価における透明性確保の取り組みとして、プレアナリシスプランの作成、研究仮説登録、再現研究についての紹介とともに、国際開発領域における透明性確保の取り組みなどの動向が紹介された。

第2報告者の岩崎久美子氏（国立教育政策研究所）からは、教育領域における政策・実践でのエビデンス活用の課題についての報告がなされた。研究者と政策立案者を対峙させた「二つのコミュニティ論（Caplan, 1979）」のレビューとともに、研究者が目指す精緻な科学的手順を経た研究成果が社会で活用されるには、そのための戦略が必要であるとの指摘がなされた。

第3報告者の津富宏氏（静岡県立大学）からは、社会的投資の観点から非営利活動を評価するツールであるSROI（Social Return on Investment）について概観の後、SROIに対する批判があることを紹介のうえで、二つの事例をもとにSROIを含む社会的投資を建設的に用いるための環境構築についての提案がなされた。

これら報告に対して多数の質問と意見があり、特にエビデンスの産出後の課題についての議論が始まったことそれ自体が特筆に値する好ましい進展であることが会場で共有された。今後もこうした介入との因果関係を科学的推論のうえに評価する研究が深まる議論を継続することの重要性を改めて共有し閉会した。なお、その後の有志らの協議により今回の報告を発展させ、第三番目のEBPの特集号を刊行する方向で合意がなされたことも併せて報告したい。

自由論題セッション報告

自由論題セッション１「参加型評価」

座長　林 薫（文教大学）

第一報告の和嶋雄一郎会員他による「学生経験調査を利用した大学生の能力に関わる分析―企業が重視する能力と大学生の主観的な能力評価の比較―」は、研究大学を中心に米国で開発された、学生経験調査（SERU：Student Experience in the Research University）の主観能力測定項目を利用し、企業がそれらの能力を採用時にどの程度重視するか（重視能力）、大学生がそれらの能力をどの程度有しているか（大学生能力イメージ）を質問紙によって調査し、「リーダーシップ」や「明瞭かつ効果的に話す能力」などの項目で、企業の重視能力と大学生能力イメージに大きな隔たりがあることなどを明らかにした。質疑応答では米国で開発された手法がどの程度日本の社会に適用できるかというコンテクスト依存性の問題や、SERUという評価手法の評価を行うべきではないかなどの点が議論された。

第二報告の新藤健太会員他による「効果的プログラムモデル形成のため実践家参画型評価アプローチ法

の開発（その7）：効果モデル改善ステップにおける「全国試行評価調査（1年間の提示版プログラム試行調査）」の位置とその検証 」では、「障害者以降支援プログラム」の事例に基づき評価と考察を行った。その結果プロセス評価、アウトカム評価を通した効果的援助要素（critical components）の検討、フィデリティ尺度（fidelity scales）、アウトカム指標など評価ツールの検討が効果的プログラムモデルの改善に貢献することが確認されたことが報告された。データに基づいた実証的研究であり、手法等について議論がなされた。

第三報告の佐々木織恵会員による「フランスの学校外部評価―目標管理型評価と参加型評価、2つの評価手法の意義と可能性についての考察―」は、近年フランスで取り入れられてきた学校外部評価の二つの異なるアプローチ（NPM の文脈で出てきた目標管理の手法と参加型評価の手法）を比較しつつ、参加型評価は評価者と被評価者の関係を対等にし、中央集権的で階層な教育行政システムを乗り越えようとする取組みであることが報告された。フランスの評価事例は学会では紹介されることが少なく、事例紹介も含め関心を集めた。

本セッションでは参加型評価の多様な在り方と評価の論点が議論され有意義なセッションとなった。

自由論題セッション２「評価の枠組み」

座長　佐々木 亮（国際開発センター）

当自由論題セッションでは、様々な評価の枠組に関して発表が行われ活発な議論が展開された。

研究報告1「イギリスにおける研究力評価の動向：REF2014のImpact評価を中心に」（藤井翔太）では、REF2014（Research Excellent Framework 2014）へのインパクト評価の導入がイギリスの大学にいかなる影響を与えたか、そしてREF2014のOutputsスコアとImpactスコアの相関係数の計算、さらに事例として取り上げたブリストル大学の対応に関する現地調査に基づいた考察が報告された。

研究報告2「環境配慮行動誘発に向けた効果的アプローチの検証　モザンビークにおけるランダム化比較試験の事例」（青柳恵太郎、細野智之）では、国際協力機構が実施している「マプト市における持続可能な3R活動推進プロジェクト」において実施した環境配慮行動に関するランダム化比較試験によって、3種類の介入の有効性が確認されたことが報告された。

研究報告3「セオリー評価における「理論」とは何か」（永井隆夫）では、「理論」とは社会問題とその原因並びに政策介入の目的と手段についての仮説であることが示され、仮説構築のための推論が論理学者パースのアブダクションという蓋然的推論であることを明らかにし、そのために仮説検証が不可欠であることが報告された。

研究報告4「学校評価計画のための論理枠の系統的作成方法についての一考察～ ‘Rainbow Framework’の紹介とその活用方法～」（橋本昭彦）では、国際的な評価実践支援グループである「ベターエヴァリュエーション」が開発した「レインボー・フレームワーク」のしくみの概要や枠組みの内容が紹介され、学校評価の計画作りへの援用の可能性について報告が行われた。

これらの報告に対して会場からは多数の質問・意見が出され、活発な議論が展開され、評価の枠組構築に関する新たな展開が感じられるセッションとなった。

自由論題セッション３「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、独立財政機関の活動内容の国際比較、知事の任期中の政策の自己評価、自治体の行政評価制度の改善、米国オバマ政権の実証重視とCBOの貢献という、

広義「政策評価」を様々な角度から論ずる4本の報告が行われた。

最初の報告である上野宏氏（統計研究会／国際開発センター）の「独立財政機関（IFIs）の活動内容とその評価：改訂版」は、第15回全国大会における報告を改訂する形で、近年先進諸国で設置が進む独立財政機関の中で代表的な8カ国の8機関を取り上げ、活動内容・独立性・組織体制などの観点から比較の上、IFIが成功するための条件を考察、米国議会予算局（CBO）型が望ましいとし、日本におけるIFI導入に向けての示唆を得ようとするものである。質疑では勧告等の強制力や米国地方政府における類似機能・機関に関する質問などが提出された。

次に、本田正美氏（東京大学大学院情報学環交流研究員）の「知事による政策の自己評価の分析」が報告された。本報告は、全国の都道府県知事が任期末に議会で開陳した任期中の政策の自己評価を、会議録の探索によって分類及び分析し、成果について集中的に質疑する会議の開催可能性など首長による政策の評価のあり方を問うものである。会場からは、行政による政策の自己評価や知事マニフェストの様々な手段による評価との関係に関するコメントなどがなされた。

3番目の報告は、高千穂安長氏（ノースアジア大学）の「自治体行政評価の改善研究」であった。本報告は、市以上の規模の自治体の行政評価担当者を対象に、筆者が経営学の観点から行ったアンケート結果に基づき、評価担当者のやる気、第三者委員会の望ましい姿、住民との良好なコミュニケーションという3つの改善をもたらす要因として、評価結果のフィードバックと動機づけ、委員会の市民感覚、評価の負担軽減等をそれぞれ抽出・指摘した。自治体行政評価の課題についてはこれまで様々な調査や議論が行われており、それらを踏まえた意見交換が会場の参加者との間でなされた。

最後の報告である、上野真城子氏（アジア都市コミュニティー研究センター）の「オバマ政権の社会政策における挑戦とCBO40年の意味するもの」は、米国オバマ政権の政策形成におけるRCTを用いた評価など実証重視の動向を紹介し、その背景として今年設立40年を迎える議会予算局（CBO）の貢献を指摘するとともに、日本にもそのような政策インフラストラクチャーが不可欠であると述べている。会場からは、米国の業績測定型評価との関係やRCTの適用政策分野などを巡る質問がなされ、また意見交換がなされた。

本セッションの報告テーマは広義「政策評価」を巡る研究の多様性を示唆するものであったが、何れもその重要性は明らかであり各研究の今後の進展が期待されるところである。

日本評価学会第 16 回全国大会
「国際評価年：評価のイノベーション」
開催の報告とお礼

2015年12月12日-13日開催（於：JICA沖縄国際センター）の日本評価学会第16回全国大会には108名の方々にご出席いただきました。誠にありがとうございました。会員各位の日頃の研究や実践活動の報告を基に、評価研究者、実務者の間の経験、情報、知識を共有化する場として、皆様にとって有益な機会となりましたら幸甚に存じます。多くの方のご尽力により本大会を開催することができましたこと、心より感謝申し上げます。今後とも当学会の活動にご高配賜りますようどうぞ宜しくお願い申し上げます。

実行委員長　　西出順郎（岩手県立大学）
プログラム委員長　牟田博光（東京工業大学）

第1日目：2015年12月12日（土）

10：00 - 10：30	受　付		
午前の部 10：30 - 12：30	**自由論題1** 社会開発評価 （西野桂子） 第1会議室	**自由論題2** 国際事業評価 （關谷武司） 第2会議室	**自由論題3** 行政機関の評価 （上野宏） 第3会議室
12：30 - 13：30	お昼休み／理事会		
午後の部Ⅰ 13：30 - 15：30	**共通論題1** ～国際セッション～ 評価の独立性、有効性のためのガバナンス （湊直信） 第1会議室	**共通論題2** 小規模自治体の行政評価 （窪田好男） 第2会議室	**自由論題4** 評価手法 （和田義郎） 第3会議室
午後の部Ⅱ 15：45 - 17：45	Symposium 国際評価年：評価のイノベーション （牟田博光） 第2会議室・第3会議室		
18：00 - 18：30	総会		
18：45 - 20：45	懇親会		

第2日目：2015年12月13日（日）

9：00 - 9：30	受　付		
午前の部 9：30 - 11：30	**共通論題3** 評価の教育 －アートかサイエンスか、それともエンジニアリングか （山谷清志） 第1会議室	**共通論題4** 科学的根拠に基づく政策形成 （青柳恵太郎） 第2会議室	**自由論題5** 教育評価 （齊藤貴浩） 第3会議室
11：30 - 12：30	お昼休み／編集委員会		
午後の部Ⅰ 12：30 - 14：30	**共通論題5** 新しい大学評価を求めて：大学内部の視点から ～評価学のフロンティアとしての「大学」～ （齊藤貴浩） 第1会議室	**共通論題6** サスティナビリティーをどう評価するか（ラウンドテーブル） （林薫） 第2会議室	**共通論題7** 行政評価局調査の現状 （村田誠英） 第3会議室
午後の部Ⅱ 14：45 - 16：45	**共通論題8** 新しい大学評価を求めて：大学外部の視点から ～評価学のフロンティアとしての「大学」～ （齊藤貴浩） 第1会議室	**共通論題9** 研究倫理と評価 （正木朋也） 第2会議室	

（　　）は座長です。

日本評価学会第 16 回全国大会
「国際評価年：評価のイノベーション」プログラム詳細

12 月 12 日（土）受付 10:00 - 10:30

12 月 12 日（土）午前の部 10:30 - 12:30

自由論題 1　社会開発評価

第 1 会議室

座長　西野　桂子　関西学院大学

J1-1	SROIを活用したソーシャルビジネスの社会的効果の検証～フィリピン小規模農民向け農機レンタル事業を事例として～	○	皆元　理恵	ジーエルエム・インスティチュート
			西野　桂子	関西学院大学
J1-2	（発表辞退）		（発表辞退）	（発表辞退）
J1-3	投資家によるプログラム評価－SIB（ソーシャル・インパクト・ボンド）の効用と課題		北大路　信郷	明治大学
J1-4	プロフェッションの倫理教育：監査人の倫理教育からの示唆		小林　信行	OPMAC

自由論題 2　国際事業評価

第 2 会議室

座長　關谷　武司　関西学院大学

J2-1	個々の子どもたちの修学パターンによる基礎教育評価	○	關谷　武司	関西学院大学
			江嵜　那留穂	関西学院大学（大学院生）
J2-2	気候変動適応における事業評価の課題		池田　まりこ	京都大学大学院（博士後期課程）
J2-3	アジアの紛争経験国を対象とした平和構築・紛争予防教育プログラムの評価	○	池田　満	南山大学
			福田　彩	東京外国語大学
			宮城　徹	東京外国語大学

自由論題 3　行政機関の評価

第 3 会議室

座長　上野　宏　統計研究会／国際開発センター

J3-1	テキストマイニングによる独立行政法人の中期目標の分析		森田　弥生	筑波大学
J3-2	協働型ネットワークにおける業績情報の学習のための活用：内部評価と外部評価による業績情報の相違についての試論		中嶋　学	ニューヨーク州立大学（博士課程）
J3-3	独立財政機関（IFI）の設立の事前評価基準と実績の事後評価基準		上野　宏	統計研究会／国際開発センター
J3-4	社会変革プログラムの評価の枠組み		上野　真城子	アジア都市コミュニティー研究センター

12 月 12 日（土）お昼休み／理事会 12:30 - 13:30

12 月 12 日（土）午後の部Ⅰ 13:30 - 15:30

共通論題 1 ～国際セッション～ 評価の独立性、有効性のためのガバナンス

第 1 会議室

座長　湊　直信　国際大学
コメンテータ　佐々木　亮　国際開発センター
上野　宏　統計研究会/国際開発センター

K1-1	Governance for Independent and Effective Monitoring and Evaluation（M&E）		廣野　良吉	成蹊大学
K1-2	Evaluation Governance and its link to Credibility, and Utility		Indran A. Naidoo	UNDP
K1-3	Role of VoPEs for Promoting Sustainable Evaluation: How Can APEA Make a Difference ？		Arunaselam Rasappan	Center for Development & Research in Evaluation, APEA

共通論題 2　小規模自治体の行政評価

第 2 会議室

座長　窪田　好男　京都府立大学

K2-1	地方公共団体における行政評価の現状―予算と評価、評価疲れ・評価廃止―		小島　卓弥	EYアドバイザリー
K2-2	地方創生における評価の位置づけ		内田　真	ジャパンインターナショナル総合研究所
K2-3	沖縄における行政評価の特殊性―沖縄振興予算の事後評価―		橋本　圭多	同志社大学

自由論題 4　評価手法

第 3 会議室

座長　和田　義郎　国際協力機構

J4-1	地方議会議員の活動の評価と検証		本田　正美	島根大学
J4-2	学校のエンパワーメントを導く評価キャパシティの検討		池田　琴恵	東京福祉大学
J4-3	規制を評価する仕組みについての一考察		村上　裕一	北海道大学
J4-4	大学教育におけるルーブリックの実践とその評価方法論的な含意		西村　君平	弘前大学

12 月 12 日（土）午後の部Ⅱ 15:45 - 17:45

Symposium 国際評価年：評価のイノベーション

第 2 会議室・第 3 会議室

座長　牟田　博光　東京工業大学

Symposium	International Evaluation Year: SDGs and Evaluation		Indran A. Naidoo	UNDP
	ADBと評価		冨永　二郎	ADB
	Panelist		齊藤　貴浩	大阪大学
			正木　朋也	国際協力機構／東京大学
			林　薫	文教大学
			山谷　清志	同志社大学
			和田　義郎	国際協力機構

12 月 12 日（土）総会 18:00 - 18:30

12 月 12 日（土）懇親会 18:45 - 20:45

12 月 13 日（日）午前の部 9:30 - 11:30

共通論題 3　評価の教育ーアートかサイエンスか、それともエンジニアリングか　第 1 会議室

座長　山谷　清志　同志社大学

K3-1	「国際的視点の評価」―セオリー評価とログフレーム，異文化理解を中心とした大学院科目―		佐藤　由利子	東京工業大学
K3-2	評価のディシプリンを再考する～公共政策大学院の講義経験から～		源　由理子	明治大学
K3-3	政策評価の教育と模擬評価の活用		窪田　好男	京都府立大学
K3-4	日本における政策評価教育の現状と課題　大学院講義20年、学部教育10年の反省		山谷　清志	同志社大学

共通論題 4　科学的根拠に基づく政策形成　第 2 会議室

座長　青柳　恵太郎　国際協力機構／東京大学

K4-1	JICAによるインパクト評価の取り組み：エビデンスに基づく事業実施		浅岡　浩章	国際協力機構
K4-2	フィリピンにおけるコミュニティ参加型インフラ整備事業のインパクト評価		倉田　正充	上智大学
K4-3	コミュニティ参加型介入によるソーシャルキャピタル蓄積―クラスターランダム化比較試験とフィールドラボ実験による実証分析―		青柳　恵太郎	国際協力機構／東京大学
K4-4	日本におけるエビデンスに基づく教育のあり方～WWC·EPPIと国内の現状から～		森　俊郎	岐阜県公立中学校／ロンドン大学（IOE）

自由論題 5　教育評価　第 3 会議室

座長　齊藤　貴浩　大阪大学

J5-1	大学認証評価における大学図書館評価の検討		高池　宣彦	筑波大学大学院
J5-2	企業調査・入学時調査・卒業時調査を利用した大学生の能力に関する分析 ―企業が重視する能力と大学生の主観的能力評価とのギャップ―	○	岡嶋　裕子	大阪大学
			和嶋　雄一郎	大阪大学
			中村　将大	大阪大学
			齊藤　貴浩	大阪大学
J5-3	大学図書館におけるスタディアシスタント（LSA）の効果に関する研究	○	渡邉　美紀	首都大学東京
			島林　智香子	首都大学東京
J5-4	SERUによる学生経験と大学における教育成果との関係	○	和嶋　雄一郎	大阪大学
			齊藤　貴浩	大阪大学

12 月 13 日（日）お昼休み／編集委員会 11:30 - 12:30

12 月 13 日（日）午後の部Ⅰ 12:30 - 14:30

共通論題 5　新しい大学評価を求めて：大学内部の視点から～評価学のフロンティアとしての「大学」～　第 1 会議室

座長　齊藤　貴浩　大阪大学
コメンテータ　田中　弥生　大学評価・学位授与機構

K5-1	大学内部に求められる評価とIR		浅野　茂	山形大学
K5-2	研究者の業績評価の現状と課題に関する研究		高谷　徹	三菱総合研究所
			林　隆之	大学評価・学位授与機構
K5-3	大学経営の効率性を高める方策の検討		山崎　その	京都外国語大学
			伊多波　良雄	同志社大学
			宮嶋　恒二	京都学園大学
K5-4	OIST、科学技術大学のあり方を模索する大学		大竹　茂行	沖縄科学技術大学院大学

共通論題 6　サステイナビリティーをどう評価するか（ラウンドテーブル）　第 2 会議室

座長　林　薫　文教大学

K6-1	趣旨説明と問題提起1　持続可能性のマクロ指標 Energy for All（第7目標）		林　薫	文教大学
K6-2	問題提起2　持続的、包括的かつ持続可能な経済成長（第8目標）、持続可能な消費と生産のパターン（第12目標）		長尾　眞文	国連大学サステイナビリティ高等研究所
K6-3	問題提起3　包括的かつ公平で質の高い教育の提供（第4目標）、グローバル・パートナーシップの強化（第17目標）		米原　あき	東洋大学
K6-4	問題提起4（研究報告）　サステナビリティ評価の手法に関する考察		加藤　郁夫	国際開発センター

共通論題 7　行政評価局調査の現状　第 3 会議室

座長　村田　誠英　総務省
コメンテータ　西出　順郎　岩手県立大学
南島　和久　神戸学院大学

K7-1	行政評価・監視（全国計画調査）「事例1：国の債権管理等」		稲垣　好展	総務省
K7-2	行政評価・監視（全国計画調査）「事例2：温室効果ガスの排出削減に係る国の補助事業」		竹中　一人	総務省
K7-3	現地における行政評価局調査の取組―地域計画調査「事例：高齢者等の交通手段の確保対策に関する実態調査」		山田　昌二	総務省

12 月 13 日（日）午後の部Ⅱ 14:45 - 16:45

共通論題 8　新しい大学評価を求めて：大学外部の視点から ～評価学のフロンティアとしての「大学」～　第 1 会議室

座長　齊藤　貴浩　大阪大学
コメンテータ　西出　順郎　岩手県立大学

K8-1	教育・研究水準を評価するための分野別基準の模索：学系別「参考例」の策定		林　隆之	大学評価・学位授与機構
K8-2	博士人材データベースの構築による人材政策への貢献	○	篠田　裕美	文部科学省 科学技術・学術政策研究所
			松澤　孝明	文部科学省 科学技術・学術政策研究所
K8-3	大学評価と市場化･社会的インパクト―大学ランキングの最新動向から―		藤井　翔太	大阪大学
K8-4	The Future of Research Metrics & its Use for Research Evaluations		Anders Karlsson	Elsevier

共通論題 9　研究倫理と評価　第 2 会議室

座長　正木　朋也　国際協力機構／東京大学

K9-1	評価 における研究倫理の諸問題		正木　朋也	国際協力機構／東京大学
K9-2	国際協力プロジェクトと研究･調査倫理:EBP実現に向けた現状課題		青柳　恵太郎	国際協力機構／東京大学
K9-3	研究不正への対応と責任ある研究活動の推進:評価の観点から		中村　征樹	大阪大学

共通論題セッション報告

共通論題セッション１「～国際セッション～評価の独立性、有効性のためのガバナンス」

座長　湊 直信（国際大学）
コメンテータ　佐々木 亮（国際開発センター）
上野 宏（統計研究会／国際開発センター）

国際セッションは英語で行い、海外の評価専門家と日本評価学会会員等の参加者が地域、国家、国際レベルにおいて評価に関する様々な課題を論じることを目的とした。特にAPEAのメンバーや国際機関からの参加者を含めた、グローバルな視点から、評価の独立性、有効性のためのガバナンスのあり方について意見交換を行った。本セッションはAPEA（Asia-Pacific Evaluation Association）との交流を促進する活動としても位置付けられており、30名以上の参加者があり、活発な議論が展開された。

第1報告者の廣野良吉氏（成蹊大学）からは、Governance for Independent and Effective Monitoring and Evaluation（M&E）と題して、モニタリング・評価の地域的、国家的、国際的な様々なレベルにおける独立性と有効性を達成するための方法についての分析と提言が行われた。第2報告者のインドラ・ナイド氏（国連開発計画）からは、Evaluation Governance and its link to Credibility, and Utilityと題し、UNDPの経験から、独立評価部門を焦点に、評価の目的、権限の分散、評価政策等について説明し、独立性重視へのグローバルな動きについて論じた。第3報告者のアルナセラム・ラサパン氏（Center for Development & Research in Evaluation, APEA） はRole of VoPEs for Promoting Sustainable Evaluation: How Can APEA Make A Difference?と題して、開発途上国の視点から評価能力強化の主要な要素、地域の評価専門家の組織の役割について報告した。更に、APEAの背景、役割、当期活動計画（意識改革、調査研究、評価能力強化、体制づくり）について説明した。

以上の報告の後、佐々木亮氏（国際開発センター）、上野宏（国際開発センター）から質問とコメントが寄せられ、それぞれの発表者が応えた。さらに、聴衆から多くの質問、コメントが寄せられ、大変に活発な議論が展開された。

日本からはいわゆる政策評価法、武蔵野市の事例なども紹介され、援助側、被援助側から様々な視点で評価の独立性と有効性を達成するための方策が論じられた。

共通論題セッション２「小規模自治体の行政評価」

座長　窪田 好男（京都府立大学）

自治体評価の職員負担がささやかれる中、小規模自治体でどのように評価を普及させていくかは日本評価学会にとって重要な課題である。本セッションの目的は自治体評価のうち、小規模な市町村で導入されている行政評価に焦点を合わせ、その現状を確認し、今後の展望を得ようとした。

本セッションの主要な論点は3つである。

第1の論点は小規模自治体における行政評価の導入状況の現状と今後の展望である。自治体評価の制度化の状況については総務省が全国的な調査を行っており、都道府県・市町村を合わせた約1800自治体のうち、約1000の自治体が制度化しているとされている。都道府県や大規模な市と比較して、小規模自治体の導入は遅れているとされる。その状況や、行政評価の職員負担とも関わる原因、今後の展望は重要な論点

であろう。

第2の論点は地方創生の総合戦略の評価への対応である。現在、国は地方創生に注力し、内閣官房にまち・ひと・しごと創生本部を設置し、人口の「長期ビジョン」と「総合戦略」を策定するとともに、地方自治体にも「長期ビジョン」と「総合戦略」の策定を求めている。その際の策定の6つのポイントの1つとして「明確な目標とKPI（重要業績評価指標）を設定し、PDCAサイクルによる効果検証・改善」をあげている。これにより、行政評価を制度化していない自治体も業績測定を手法とする評価制度の導入が求められることになった。

この地方創生「総合戦略」の業績測定が小規模自治体にどのように導入されるのか、適切に運用されるのか、導入され適切に運用されるための条件は何か、導入され適切に運用されるためにどのようなことがなされるべきか、論点は多数ある。

第3の論点は小規模な自治体における行政評価の地域特性である。行政評価の地域特性という論点は新たな論点であり従来は見られなかったものである。本セッションでは沖縄振興予算の事後評価を取り上げる。沖縄振興予算は沖縄の成長を目的とする国の政策に基づくものである。沖縄県内の小規模自治体の行政評価の導入率は全国平均を下回っている。一方で、沖縄振興予算については小規模自治体を含む沖縄県内の全自治体が評価に取り組んでいる。この事例をもとに行政評価の地域特性について議論を重ねるきっかけとしたい。また、この論点は第2の論点とも密接に関係している。

冒頭に司会者（窪田）より趣旨説明および論点の整理を行った。

第1報告者の小島卓弥氏（EYアドバイザリー（株））からは、「地方公共団体における行政評価の現状 ―予算と評価、評価疲れ・評価廃止―」と題し、最新の状況について報告があった。小島報告も踏まえ、第2報告者の内田真（(株)ジャパンインターナショナル総合研究所）からは「地方創生における評価の位置づけ」、第3報告者の橋本圭多氏（同志社大学）からは「沖縄における行政評価の特殊性 ―沖縄振興予算の事後評価―」について報告が行われ、討論者（窪田）からのコメント後、会場との質疑応答、意見交換が行われた。

共通論題セッション３「評価の教育―アートかサイエンスか、それともエンジニアリングか」

座長　山谷 清志（同志社大学）

政策評価、ODA評価をはじめとして日本の高等教育機関では20年ほど評価教育を行ってきた。しかし実際にその教育内容には標準的なカリキュラムはなく、共通して使用できる教科書もない。早い時期には社会工学や計量経済学・計量社会学的なアプローチが志向され、その後ODA政策、政治学や行政学の分野では教員がそれぞれ工夫して編み出してきた職人芸的なアート（スキルとテクニック）が増えてきた。もちろんエビデンスにもとづく評価を重視する科学的な接近方法も大事にされる。こうして評価教育と言っても、実にさまざまなバリエーションが存在するようになった。問題なのは、評価学会に所属する大学教員が、他の教員がどのような講義をしているかほとんど知らないことである。このセッションでは標準的な教育カリキュラム作成の可能性と評価教育の質的向上をめざし、その準備作業として大規模の学部教育、少人数の学部教育、そして大学院の講義科目、あるいは大学院の演習科目それぞれについて、どのような教育をしているのか実情を報告してもらった。

第1報告者の佐藤由利子氏（東京工業大学）からは、「「国際的視点の評価」－セオリー評価とログフレーム、異文化理解を中心とした大学院科目－」と題して、理科系の留学生が多い大学院において評価文化の違いの理解を進め、それを前提に評価スキルの修得をめざす教育方法について報告していただいた。評価スキルとはログフレーム作成、ステークホルダー文責など多岐にわたるが、いずれも国際的な評価教育

のスタンダードである。佐藤氏の教育の特徴は、評価手法の習得後に、身近な事象への応用など、実践的な教育を行っているところである。

第2報告者の源由理子氏（明治大学）は、「評価のディシプリンを再考する～公共政策大学院の講義経験から～」と題して、評価学とは何かを改めて考えつつ、標準的な評価カリキュラムを展望された。その際、ディシプリンとしての評価学は、アメリカと日本とでは評価技能や評価技法の背景にある知識体系と、その運用方法が違うことから、日本の社会背景（評価人材市場の限界）にマッチしたシラバスの必要性を考えられていた。

第3報告者の窪田好男氏（京都府立大学）は、公立大学という特徴、すなわち少ない学生定員から政策評価講義がかなり小規模であり、その小規模を活かしたシラバスづくりをされていると報告された。ケースメソッド、ロールプレイ、アクティブラーニングである。「地域公共政策士」（京都府と京都市内の大学が連携して認証する資格）の対象科目になっているこの講義は、政策評価教育のあり方を考える際の、一つのモデルとなっている。

第4報告者、山谷清志は「日本における政策評価教育の現状と課題　大学院講義20年、学部教育10年の反省」大規模大学における政策評価教育の課題について報告した。数百人が履修するため、座学にならざるを得ず、また応用社会科学的な特徴を持つ政策評価論講義も、かつての大学のように一般教育が存在しないため、政治学、経済学、法律学、社会学を履修していない学生が教室に多い。政策評価教育の前提が出来ていないのである。大学院ではそうした大規模講義の難点はないが、こんどは政策評価の前提知識を持たない大学院生が、実践的な問題関心から突然ODA評価、防衛政策、外交政策の評価を考えるという課題がある。

このセッションは教育方法をめぐる討論であったため、大学関係者にとっては大変参考になる内容であったと自負している。近年教育方法の諸問題について文部科学省をはじめとしてさまざまな議論、改善案が提示される中、現状に満足していない大学教員にとってこのセッションはいろいろな意味で参考になったと思われる。

共通論題セッション４「科学的根拠に基づく政策形成」

座長　青柳 恵太郎（国際協力機構／東京大学）

日本評価学会の社会実験分科会がEBP（Evidence-based Practice）の普及啓発活動を開始した2005年から10年が経過した。この間、関係各位の努力により様々な領域において、実施プロジェクトや施策の因果効果を評価するための枠組みについて理解が深まり、実践事例の蓄積も着実になされてきた。本セッションでは、科学的根拠に基づく政策形成（Evidence Based Policy Making）を論じた近年の研究動向を踏まえ、日本国内におけるEBPの方向性を見据えることを目的に、国内の国際開発及び教育領域から4つの事例報告が行われた。

冒頭、座長より現在のEBP、及びその根幹をなすインパクト評価を論じた研究動向として、①評価デザインの工夫・実施環境の多様化、②アウトカム測定方法の精緻化、③効果発現に到るメカニズム解明の取り組み、そして④インパクト評価の結果活用に関する研究の進展の4つが指摘された。

第一報告の浅岡浩章氏（国際協力機構）「JICAによるインパクト評価の取り組み：エビデンスに基づく事業実施」では、近年EBPに向けた取り組みが大きく進展している国際協力機構におけるインパクト評価の実施状況と結果活用に関する課題が報告された。結果活用を促すためには研究目的のKnowledge-focused evaluationsではなく、政策決定の判断にかかるDecision-focused evaluationを実施していくことが重要であるとの指摘がなされた。続く、倉田正充氏（上智大学）、及び青柳恵太郎氏（国際協力機構／東京大学）か

らは、それぞれ「フィリピンにおけるコミュニティ参加型インフラ整備事業のインパクト評価」、「コミュニティ参加型介入によるソーシャルキャピタル蓄積－クラスターランダム化比較試験とフィールドラボ実験による実証分析－」が報告された。いずれも国際協力機構が行ったインパクト評価の事例報告であり、前者は紛争復興期という厳しい環境下において実施されるプロジェクトを対象としたエビデンスをつくるための試みであり、後者は認知的ソーシャルキャピタルという目に見えない（数量データとして観測することが難しい）性質のアウトカムを心理学・行動経済学に基づくフィールドラボ実験と呼ばれる手法で計測したという特徴を有していた。最後に、森俊郎氏（岐阜県公立中学校／ロンドン大学（IOE））から、「日本におけるエビデンスに基づく教育のあり方 ～ WWC·EPPIと国内の現状から～」という演題の下、教育分野におけるエビデンスの産出からその活用と政策形成に向けての課題について、英国EPPIセンター（The Evidence for Policy and Practice Information and Co-ordinating Centre）、アメリカWWC（What Works Clearinghouse）の現状などが報告された。

各報告、及び会場からの質問・意見を通じて、「科学的根拠に基づく政策形成」への関心が国際開発、教育領域のみならず広範な分野において高まっており、各種取組が始まっていることが確認できた。今後とも、領域横断的に「科学的根拠に基づく政策形成」に向けた議論が継続していくことを期待する。

共通論題セッション５「新しい大学評価を求めて：大学内部の視点から ～評価学のフロンティアとしての「大学」～」

座長　齊藤 貴浩（大阪大学）
討論者　田中 弥生（大学評価・学位授与機構）

教育・研究は評価が難しい活動である。その中でも大学は、教育研究活動が高度であるばかりか、多様な利害関係者とそのニーズが複雑に絡まり、「評価学」にとっては依然としてチャレンジングな対象であると言える。大学設置基準の大綱化に伴い自己評価の努力義務が課されてからおよそ四半世紀が経過し、第三者評価としての認証評価は第2サイクルの終盤に、そして国立大学法人は第2期の国立大学法人評価を迎える。しかし、未だに確固とした評価方法が確立されたとはいえない。

本セッションは、大学評価を内部と外部とに分けて議論を行ったうちの「内部の視点」のセッションである。大学は様々な利害関係者に囲まれ、また様々な視点で期待されており、内部と外部とでは大きく異なることからそれらを対比する構成とし、対となる「外部の視点」は共通論題セッション8として、本セッションの次の時間に開催された。

第1報告者の浅野茂氏からは、昨今、大学において「必要な情報を必要なときに提供し、大学の意思決定を支援する一連の活動」としてのインスティテューショナル・リサーチ（IR）への興味・関心が高まっているものの、それは長い間唱えられてきた、大学評価の基本は各大学が自ら行う自己点検・評価にあるということと本質的には違わず、いかに大学内の活動を自らの評価や外部からの評価に結びつけていくかが重要であり、その具体的な取り組みについて実例を挙げながらご紹介をいただいた。第2報告者の高谷徹氏、林隆之氏からは、本来組織マネジメントにとって重要な人事評価が大学でどのように行われているか、すなわち研究者の業績評価の現状と課題について、文部科学省の委託調査の結果を基に明らかにした結果、研究者の業績評価制度は特に国立大学で導入が進んでいるものの、それは制度として求められているからにすぎず、依然として評価結果の効果的な活用、例えばモチベーションや活動計画への反映等には十分な取り組みが行われていないことが示された。第3報告者の山﨑その氏、伊多波良雄氏、宮嶋恒二氏からは、私学の大学マネジメントにとっては必須であり、昨今国立大学の在り方についても議論されている大学経営の効率性について、大学の活動のインプットとアウトプットをDEA（包絡分析法）によって分

析・評価する方法論と、大学を属性や指標によってグループに分けた分析結果や個別大学の改善案の例などを提示していただき、これが各大学の効率性の評価の基礎となりうることが示された。最後に、第4報告者の大竹茂行氏からは、大学の内部の評価の実例として、沖縄科学技術大学院大学の事例についてご説明をいただいた。同大学は日本政府の構想に基づき政策的に設置された大学であり、真に国際的で、教育と学際的研究において世界の最高水準の大学となり、沖縄の持続的経済発展に貢献することを目的として掲げており、現状分析の結果、大学は独立した存在であり得ず、外部環境としての社会や学術界との関係性を視野に入れつつ課題解決を図り、目的を達成することの必要性について紹介がなされた。

共通論題セッション６「サステイナビリティーをどう評価するか（ラウンドテーブル）」

座長　林 薫（文教大学）

2015年9月に国連総会で持続可能な開発目標（Sustainable Development Goals: SDGs）が採択された。SDGsは17目標169項目からなり、それらに対応する指標の設定作業が現在進められているが、SDGsの達成状況、さらには持続性とは何かという本質的な問いかけが評価の課題となるには必然である。本セッションはラウンドテーブル形式で、まずは評価としてアジェンダ設定のための自由な議論を行い、学会としてサステイナビリティーに関する研究分科会を立ち上げるための基礎固のために企画したものである。

最初の趣旨説明と問題提起は座長である林会員が行った。2015年9月に持続可能な開発目標（SDGs）が採択されたが、持続性そのものを測る指標が必要であること、持続性を指標化する場合には、国内総生産（GDP）のような付加価値ベースのフローの指標のみならず、枯渇性資源の使用などをストックの減少としてとらえて、指標化していく必要があることがSDGs期間中の大きな課題であることなどが説明された。

第2の問題提起は長尾眞文会員によるものである。SDGの特徴のひとつは、その多くが達成度目標と取組目標の合成的表現として構成されていることで、その評価にも多様なアプローチの存在が考えられるとし、持続可能な経済成長と生産的な雇用の促進に関わる目標8との関連で、アフリカの主要な政策関心として浮上しつつある持続可能な工業化の課題に着目し、その実現を担うべき企業家人材の育成プログラム構築について問題提起した。取組目標の持続可能性の評価では、工業化の主要なステークホルダーである既存企業の指導者層の参加が不可欠であることが挙げられた。

第3の問題提起は米原あき会員によるもので、SDGsのゴール4と17に注目し、「指標」と「モデル」の観点から議題を提供した。SDGsの全体的な特徴として、各ゴールのもつ高い包括性が挙げられる。この包括性を捉えるため、人間開発指標に見られるような「複合指標」の開発が望まれるところであるが、さらに異なるゴール間にも包括性が見いだせるSDGsにおいては、指標のみではなく、複数の複合指標を活用した「モデル」によるアプローチを検討する必要性がある、また、SDGsのような複雑な開発目標の評価に際しては、「指標」による成果測定に基づく相対評価のみならず、改善に向けてのプロセス評価を重視する形成的なアプローチも重要であると考えられることなどが論じられた。

加藤郁夫会員の報告「サステナビリティ評価の手法に関する考察」はラウンドテーブルセッションの中で、通常の報告形式で行われた。欧米では1990年代に投資先の企業やプロジェクトのサステイナビリティーを考慮に入れるミューチャルファンドが急増し、国連責任投資原則（PRI）に署名済みの機関投資家の資産運用は2014年4月末で45兆米ドル超に膨らんでいるが、サステイナビリティーをどのように評価するかが課題になっており、これに対応するために試みられつつあるバランス・スコアカードの仕組みと問題点について報告がなされた。

質疑応答ではサステイナビリティーの評価手法、評価基準等についてさまざまな観点からの自由かつ活発な議論が行われた。

共通論題セッション７「行政評価局調査の現状」

座長　村田 誠英（総務省）
コメンテータ　西出 順郎（岩手県立大学）
南島 和久（神戸学院大学）

総務省行政評価局は、国民に信頼される質の高い行政を実現するための業務を行っており、その業務は行政評価局調査、政策評価の推進、行政相談に大別される。今般の発表テーマである行政評価局調査とは、総務省行政評価局が、施策や事業を担当する府省とは異なる立場から、複数府省にまたがる政策や府省に共通する制度や手法を活用する政策を横断的に評価する「統一性・総合性確保評価」及び対象を特定の政策に限定せず府省の業務の実施状況をチェックする「行政評価・監視」の総称である。また、行政評価・監視は、全国的規模で行う「全国計画調査」と地方支分部局である管区行政評価局・行政評価事務所が地域に密着した行政上の課題を取り上げる「地域計画調査」に大別される。

具体的な発表内容として、稲垣好展氏（総務省行政評価局評価監視官）は全国計画調査である①「国の債権管理等に関する行政評価・監視」を、竹中一人氏（総務省行政評価局評価監視官）は全国計画調査である②「温室効果ガスの排出削減に係る国の補助事業に関する行政評価・監視」を、山田昌二氏（総務省沖縄行政評価事務所長）は地域計画調査である③「高齢者等の交通手段の確保対策に関する実態調査」を例に、それぞれの課題や総務省が行った勧告の概要等を報告した。

コメンテータである西出順郎会員からは、①に関し平成19年に行った債権管理に関する調査と今般取り上げた平成27年の調査との関係等について、②に関し行政評価局調査と各府省が行う政策評価の関係等について、③に関し地域における課題と行政相談の関係等について質問及び指摘がなされた。同じくコメンテータである南島和久会員からは、①に関し債権管理とマイナンバーとの関係等について、②に関し政策の効果測定と外部要因の関係等について、③に関し総務大臣の勧告権限と地方分権の関係等について質問及び指摘がなされた。

行政府における評価専担組織として実務を担っている総務省行政評価局は、これまでも全国大会において行政評価局調査に関する発表を重ねている。本学会において総務省行政評価局の取組についての理解が更に深化していると考えており、今後も交流を図りながら、質の高い取組を進めていくことで国民に信頼される質の高い行政の実現に務めてまいりたい。

共通論題セッション８「新しい大学評価を求めて：大学外部の視点から ～評価学のフロンティアとしての「大学」～」

座長　齊藤 貴浩（大阪大学）
コメンテータ　西出 順郎（岩手県立大学）

本セッションでは、共通論題セッション5に引き続き、大学外部の視点から大学評価の進展を検討した。大学外部からの評価としては認証評価や国立大学法人評価が行われているが、昨今では、これらの公的評価に加えて、大学ランキングのような外部機関が行う評価に関する動きが活発になっている。例えば、大学ポートレートや、U-Multirank のように、情報公開とこれから教育を受ける人が大学を選択することを主眼に置いたツールもまた、大学評価の一環として機能しつつある。一般に、大学外の者は大学の活動を直接的に把握することはできない。また、大学で行う教育研究は高度であったり、専門的であったりするために、普通の人が判断を下すことは難しく、結果として情報公開の中での「価値付けとしての評価」の役割も必要不可欠であろうと考えられる。大学は、より一層の情報公開とともに、このような利害関係者

の「見方」や「見え方」をも想定した対応が必要とされていると言える。

第1報告者の林隆之氏からは、2016年に行われる国立大学法人の教育・研究評価における学部・研究科単位の現況分析（教育・研究の水準評価）において、評価者が優れた事項の例として参照可能な分野ごとの「参考例」の作成を行う過程を対象に、大学の自律性の尊重と外的な誘導、さらには評価可能性、比較可能性を勘案した上で、「参考例」策定自体の是非について示されるとともに、どのような議論がなされ、どのような結果が議論で得られたのかについて検討を行った経緯と今後の課題が示された。

第2報告者の篠田裕美氏、松澤孝明氏からは、博士課程修了者の属性や修了後の継続的なキャリア追跡を可能とする情報基盤として科学技術・学術政策研究所が整備する「博士人材データベース（JGRAD）」構築による人材政策への貢献」についての紹介がなされた。直接的な大学評価ではないが、大学院教育の成果である博士人材を情報として把握し、効果的な政策形成に活かそうとする事業であり、「政策のための科学」の一貫としての取組であることからもわかるように、今後の大学セクターに関する政策評価の先駆となることが期待される。

第3報告者の藤井翔太氏からは、大学評価の市場化、もしくは商業化について、THEやQS、U-Multirankや英国のREFを事例として紹介された。それらが単なるランキングから、比較可能なデータ提供者としての振る舞いをするようになっていると指摘し、これからの大学はデータの可視化や二次利用をも念頭に置いた対策をする必要があることが論じられた。最後に、Anders Karlsson氏は、学術論文データベースを大学に提供する立場から、引用情報を組み込んだ学術論文データベースがより拡張し、また高度化し、様々な分析が可能となったことで研究者や組織にとってもその可能性が拡がり、また英国のREFや、日本の国立大学法人評価でも使用されるようになり、もはや大学評価に不可欠な存在となっていることを示された。

共通論題セッション5、共通論題セッション8を通じて、本共通論題セッションで発表をいただいた課題は、各発表とコメンテーターからのコメント、および参加者間の質疑応答と意見交換を通じて、それぞれ新たな大学評価の姿として更なる実践と研究の方向性を示していることを参加者間で共有することができた。知識基盤社会をリードし、わが国の可能性を示すべきこれからの大学が何をすべきか、それとともに「評価学」として、大学にどのような貢献ができるのかを継続的に検討することが必要であると考えられる。

共通論題セッション９「研究倫理と評価」

座長　正木 朋也（国際協力機構／東京大学）

インパクト評価の普及に伴いその実施に先立ち倫理審査を行うことが必要との潮流があり、研究論文の執筆指針にも倫理審査機関からの承認を掲載要件とする旨を明記する学術誌も増えている。このような背景のもとで、事業として実施する国際協力プロジェクトにおいても研究と同レベルの倫理審査の必要性の有無についての議論が生じてきた。本セッションでは、このような動向を踏まえ、事業の評価を行う際の研究倫理(research ethics)に関わる諸問題を取り上げ、今後、どのような課題があるかを議論した。

冒頭で座長よりこのセッションの背景と趣旨説明および論点提起を行い、研究倫理に関わる議論を行う際の留意点―「被験者保護の観点からの議論」と「研究不正の問題」は明確に区別して議論すべきものであること―を強調したうえで各演者からその立場を明確にした報告がなされた。

第1報告者（座長）より、倫理的問題そのものの取り扱い難さに加えて、論争となっている問題そのものの認識共有が難しい状況が生じ得る背景などをレビューのうえで、主に「被験者保護の観点からの議論」を中心に、これら課題を解決するための方策としての倫理指針整備の歴史とその意義、また、課題解決策としてシナリオ検討と決疑論が役立つであろうことを報告した。

次いで第2報告者の青柳恵太郎氏（国際協力機構／東京大学）からは、国際開発の現場で経験した事例

をもとに具体的な問題点についての報告がなされた。インパクト評価における倫理的配慮に関わる議論を例として、国際協力プロジェクトの中で行われる社会調査（必ずしも論文投稿や研究報告を目的としないもの）に対して実施する倫理審査に関して、①対象範囲、②実施主体、③データの二次利用、④スケジュールとの兼ね合いといった4つの視点に整理し、論点を明示したうえでフロアからの意見を求めた。

第3報告者の中村征樹氏（大阪大学）からは、国内外で経験した「研究不正の問題」について、具体的事例とともに現状を俯瞰する教育的内容の報告がなされた。狭義の研究不正行為（特定不正行為: 捏造・改ざん・盗用）に限らず、二重投稿や不適切なオーサーシップ、研究記録の不適切な管理や研究成果の意義を過大視させるような統計の利用、実験方法詳細の不記載などの好ましくない研究行為が研究活動に負の影響を与えるものとして問題視されている現状を示したうえで、「研究不正の問題」への対策は「責任ある研究活動」を推進することにより質の高い研究成果を産み出していく方向へと、取り組みの重点がシフトしていることが共有された。

これら報告に対するフロアからの質疑を通じて、研究倫理の諸問題には制度や防止策に頼ることなく、自己規律として認識し対処すべきであることが再確認された。また、事例として報告のあった医学領域によらず環境分野においても「不の影響」についての議論もあり、今後も領域横断的な検討を継続することが重要である点を共有し閉会した。

自由論題セッション報告

自由論題セッション１「社会開発評価」

座長　西野 桂子（関西学院大学）

近年、インパクト評価に代表されるように、社会開発評価の定量化・数値化が進んでいる。介入から得られた効果を数字で「見える化」することにより、ステークホルダーの理解を促進する利点がある。また、評価者の倫理への関心も高まっている。本セッションは、「社会開発評価」という切り口の下、以下の3つのテーマが報告された。

まず、皆元理恵会員（特活ジーエルエム・インスティチュート）が、「社会的投資収益率（SROI）を活用したソーシャルビジネスの社会的効果」を報告した。フィリピンで実施中の農機レンタルビジネスで新規活動を計画する際、SROI値を算出して社会的効果と経済的（事業的）効果の比較検討を行った。通常のビジネスとは異なり、社会性と事業性のバランスが求められるソーシャルビジネスの計画段階において、見えない「社会的効果」を貨幣価値に換算するSROIは、事業内容の評価検討に有効なツールであるとの結論であった。

田中博会員（参加型評価ファシリテーター）の発表辞退による休憩の後、北大路信郷会員（明治大学）が「投資家によるプログラム評価―SIB（ソーシャル・インパクト・ボンド）の効用と課題」について発表した。SIBとは、特定の社会的介入プログラムの期待成果と実現可能性を投資家が評価し、プログラムに必要な資金を提供し、社会的介入を実行する仕組みである。社会的なリターンを期待する「寄付」と経済的なリターンを期待する「金融投資」の中間をなし、両方のリターンを求め、リスクを伴う投資行動を指す。英国でのSIB設計例を基に、今後日本でも、SIBを社会的介入サービスに応用することが重要であるとの見解が示された。

最後に小林信行会員（OPMAC(株)）が「プロフェッションの倫理教育：監査人の倫理教育からの示唆」を報告した。評価者の倫理に関しては、多くの議論がなされているものの、罰則など強制力を伴う制度が整備されていない。一方で罰則が伴う監査人の場合は、ケース・メソッドを活用して意思決定プロセスの習得が重視され、職業倫理の醸成がなされている。状況を読み解く能力が求められる評価者にとって、ケース・メソッドは意義のある教育手法であり、ケース・メソッドを通じ評価倫理ガイドラインの適切さを論じることは、検討に値するとの提案がなされた。

事業の社会性と経済性の比較検討、社会的および経済的なリターンを求めるSIB、社会開発評価を行う評価者の倫理教育という幅広い内容であったが、貨幣価値への変換方法、比較の意義、日本にあった制度設計、倫理研修の方法論など質疑応答や討論が活発に行われ、今後も継続的に議論がなされるべきテーマであることが確認された。

自由論題セッション２「国際事業評価」

座長　關谷 武司（関西学院大学）

本セッションでは、開発途上国の教育評価における縦断的データ活用の可能性、気候変動適応における事業評価の検証、そして、平和構築・紛争予防教育プログラムの評価について興味深い発表をいただいた。以下、簡潔に個別の発表を振り返る。

江嵜那留穂氏（関西学院大学国際学研究科）から、個々の子どもたちの修学パターン（縦断的データ）を用いた基礎教育評価の検討を報告いただいた。これまでは、各国の教育状況をみる際、就学率や退学率といった数値で評価されることが多いが、これらの数値は横断的データの集計値であり、一人一人の修学実態や子どもたちが抱えている課題を明らかにすることはできない。そこで、学校記録を丹念に精査した縦断的データから修学パターンを明らかにし、中米・ホンジュラス、サブサハラ・ザンビア、南アジア・ネパールの初等教育学校の修学状況の比較から、縦断的データを用いた評価の可能性について検討を行った。

池田まりこ氏（京都大学大学院地球環境学舎）は、気候変動対策として「緩和」とともに両輪として機能する「適応」プログラムについて、国際的組織である「適応基金」がセネガルで実施した事例を対象とし、途上国における適応プログラムのモニタリングや評価がどのように行われたのかを検証した。その結果、実施主体のガバナンスの問題、事業実施計画における問題、投入とアウトカムに対する自己評価および外部評価の明確性などの課題があげられた。

池田満氏（南山大学）は、東京外国語大学を中心に行われている平和構築・紛争予防教育プログラムを対象とし，その効果として，受講者の文化的知能（文化的多様性に対応する能力）と選択的道徳不活性（特定の状況下で道徳的判断から逸脱した意思決定を行う程度）を測定した結果について報告された。プログラムの成果としては、文化的知能の上昇は見られたものの、選択的道徳不活性についてはプログラム前後で変化が見られなかったことが示された。

これらの報告に対してフロアから質問・意見が出され、政策への提言の可能性や、外部評価としての視点、得られた数値の意義などについて議論が行われた。開発途上国や紛争国などにおける国際事業に関する評価について、今後もさらに現場に根ざした研究発表が重ねられることが期待される。

自由論題セッション３「行政機関の評価」

座長　上野 宏（統計研究会／国際開発センター）
佐々木 亮（国際開発センター）

行政機関の評価はますます重要性を増しており、その手法も様々なものが提案されている。本セッションでは、まず、森田弥生会員（筑波大学）が「テキストマイニングによる独立行政法人の中期目標の分析」を報告した。同報告では、独立行政法人通則法が一部改正され、独立行政法人制度は新制度がスタートしたことを受け、テキストマイニングを用いて中期目標を分析することにより、目標の特徴について明らかにした試みが報告された。

次に、中嶋学会員（ニューヨーク州立大学（博士課程））が、「協働型ネットワークにおける業績情報の学習のための活用：内部評価と外部評価による業績情報の相違についての試論」を報告した。同報告では、米国の「システム・オブ・ケア」を事例にして、協働型ネットワークにおける業績情報の学習のための活用に影響を与える要因、さらに、内部評価と外部評価により産出された業績情報の間で要因に相違があるかが明らかにされた。

続いて、上野宏会員（統計研究会／国際開発センター）が、「独立財政機関（IFI）の設立の事前評価基準と実績の事後評価基準」を報告した。同報告では、OECDが提言しているIFIの事前評価基準と事後評価基準を検討し、両者が整合性を保つ努力をしていること、事後評価基準にはかなりの問題点あることが明示された。

最後に、上野真城子会員（アジア都市コミュニティー研究センター）が、「社会変革プログラムの評価の枠組み」を報告した。同報告では、米国でここ10年ほどに注目されてきた近隣指標システム（Neighborhood Indicator System）を取り上げ、そのシステムが、市民の継続的内発的なコミュニティー・ビルディングの形成のダイナミズムを生み出すこと、参加によるコミュニティー変革の「運動」のツールとなることを紹介し、日本の地方創生と活性化において、極めて有効な戦略的計画手法となることが示された。

これら4本の報告に対して会場からは多数の質問・意見が出された。日本の行政機関が用いることができる新しい評価手法については、今後とも継続的に議論がなされる必要があろう。

自由論題セッション４「評価手法」

座長　和田 義郎（国際協力機構）

本セッションにおいては、地方政治、初等教育、高等教育、規制といった、公共性の高い分野における評価の手法について、議論を行った。

第1報告者の本田正美氏（島根大学）からは、「地方議会議員の活動の評価と検証」として、長崎県大村市議会の議会での質問に対する評価の試行について、その結果についての発表があった。

ついで、第2報告者の池田琴恵氏（東京福祉大学）からは、「学校のエンパワーメントを導く評価キャパシティの検討」として、小学校で行われる自己評価において、エンパワーメント評価を用いた実践についての検討結果から見出された4つの組織的評価キャパシティについての報告があった。

第3報告者の村上裕一氏（北海道大学）からは、「規制を評価する仕組みについての一考察」として、米国行政管理予算局（OMB）の情報・規制問題室(OIRA)における規制審査を他の国内・国際事例と比較して、規制インパクト評価の仕組みを検討した結果についての規制評価の合理性・独立性・網羅性の観点からの整理について報告があった。

第4報告者の西村君平氏（弘前大学）からは、「大学教育におけるルーブリックの実践とその評価方法論

的な含意」として、現在、初中等教育から導入され、大学教育において急速に普及し始めているルーブリック評価について、その評価手法の検討についての発表があった。

各報告における質疑を通じて、議会活動や規制などの公共政策、初等、高等教育において既に導入されている様々な評価、規制インパクト評価やルーブリック評価、エンパワメント評価などについての評価手法の比較、検討、提案などがなされた。教育や公共政策における評価の重要性は言うまでもなく、また、最も導入の進んでいる分野であるが、その中でも、どのような特性に重点を置くか、実際に起こっている現実をどのように評価に取り込むか、という点で、評価手法の検討がきわめて重要であることが認識された。今後は、評価手法を一層洗練するとともに、個別の事象を取り込めるより特定の評価手法の開発が議論を通じてなされることを期待するものである。

自由論題セッション５「教育評価」

座長　齊藤 貴浩（大阪大学）

教育評価のセッションでは、すべて大学の教育に関連した研究発表となった。2件が質問紙を用いた教育の成果に関する研究、2件が大学の図書館の機能に関する研究であり、うち1件は図書館が提供するサービスの成果に関する評価、もう1件は図書館自体の評価の在り方に関する研究であった。これらの研究発表を通じ、参加者との間で活発な質疑応答と意見交換が行われた。

高池宣彦会員（筑波大学大学院）「大学認証評価における大学図書館評価の検討」は、認証評価における大学図書館評価のあり方について、国際規格の図書館評価との比較によりプロセスの評価が欠落していることを指摘し、認証評価でもそのような評価を行うことで図書館評価の在り方をガイドするなど、具体的な改善策について提案した。岡嶋裕子会員、和嶋雄一郎会員、中村将大会員、齊藤貴浩会員（大阪大学）「企業調査・入学時調査・卒業時調査を利用した大学生の能力に関する分析」は、大学で身につけるべき能力について、企業が採用時に重視する程度と大学生および卒業生が有している程度について質問紙調査を行った結果、「柔軟な思考」「自由な発想」が不十分であるなど、ギャップがあることを明らかにした。渡邉美紀会員、島林智香子会員（首都大学東京）「大学図書館におけるスタディアシスタント（LSA）の効果に関する研究」は、図書館で学習をサポートするLSAの活動に関する質問紙調査の結果をもとに現状を分析し、LSAの名前は知っているもののその役割やメリットが十分に伝わっておらず、今後改善の余地があるとの事業評価の結果を紹介した。最後に和嶋雄一郎会員、齊藤貴浩会員（大阪大学）「SERUによる学生経験と大学における教育成果との関係」は、学生の大学での経験に関する調査結果と、学生の学習に関わるデータとを統合し、学生の大学での経験が学生の成績や就職活動、大学への満足度などにどのような影響を与えているのかの検討を行った結果を紹介した。

これらの研究内容を俯瞰すれば、単に教育や教育に関連する施策が上手くいったかを成果のみで評価するのではなく、成果を生むまでのプロセスを大事にするという姿勢でもあり、今後の教育評価に関する研究を厚くしていくための一つのアプローチと言えるのではないかと考えられる。ただし、プロセスの研究は少なからず実態把握で終わってしまうこともあり、評価論の立場からは投入や成果との関係や今後の方策、教訓へのさらなる進展が求められるはずで、次回以降も多くの研究報告と意見交換が行われることを期待したい。

シンポジウム報告

「国際評価年：評価のイノベーション」

座長　牟田 博光（東京工業大学）
キーノート・スピーカー　Indran A. Naidoo（UNDP）
討論者　冨永 二郎（ADB）
齊藤 貴浩（大阪大学）
正木 朋也（国際協力機構／東京大学）
林 薫（文教大学）
山谷 清志（同志社大学）
和田 義郎（国際協力機構）

2015年は国際連合の評価グループから評価と証拠に基づく政策形成を提唱する「国際評価年」と指定され、2014年の国連総会でも、国単位での評価能力の向上について決議が行われている。2015年は、ミレニアム開発目標の最終年であり、2016年度以降のSDGsも見越して、評価への要請が高まることも考えてのことだろう。わが国でも、2015年7月に参議院本会議で政策評価制度に関する決議が行われている。

本シンポジウムでは、キーノート・スピーカーのNaidoo氏からは“International Evaluation Year: SDGs and Evaluation”というテーマでお話しを頂いた。2030年を目標年次としたSDGsはMDGsとは異なり、単に発展途上国だけの目標ではなく、環境問題に象徴されるように、地球上のすべての国々の目標となるものであり、SDGsの実現を目指して、それぞれの国でNational Evaluation Capacityを高めて、評価に力を入れる必要があり、評価の仕組み作りや評価者の訓練なども必要になる、との強いメッセージが報告された。

第1報告者の冨永氏からはADBの業務評価のイノベーションの話があった。評価の独立性を行動上、組織上、外部からの影響の排除、利益相反の排除の観点から確保し、公正性、信憑性、有用性、パートナーシップ、ラーニングの奨励の原則を守ることが重要で、評価の影響力を高めるために、戦略的評価計画の策定として、適切な時期に適切なトピックを選択すること、評価の質の向上として、分析結果の有用性の向上、具体的かつ戦略性に富んだ提言の作成、各種評価手法を取り入れること、独立性とエンゲージメントとして、独立性の確保と業務部局との関係の維持、評価結果の公表と積極的な普及を行っているとの報告があった。

第2報告者の齊藤氏からは、現在の大学評価についての3つの動きについて発表があった。予算を配分する政府から、政策誘導としての評価のみならず、事前評価に政策課題を評価項目として含めることで改革を進めようとする動きがある。費用を負担する社会を念頭に置いた大学ランキングや、大学ポートレイト、博士人材DBなどの情報公開・情報共有の動きがあり、それを政府が活用する動きがある。機関単位の大学評価から、大学の学問分野、機能に着目した評価や、大学自らの自己評価の充実に力点が移りつつある。これらの動きは、大学・政府・社会という3つの主体のバランスが変化していることを意味しており、新たな大学評価システムの模索が必要であるとの報告がなされた。

第3報告者の正木氏からはインパクト評価は2000年以降国際的に急速に広い分野・領域で実施されている現状が示され、同時に、被験者保護の観点から倫理的事案をどのように取り扱うべきかの議論が生じている旨の報告がなされた。2012年に日本評価学会で策定した評価倫理ガイドラインも参照しつつ、今後の事業評価において、どのような場合に倫理審査を必要とするかなど、学会として一定の方向性を示せるよう検討することが必要ではないか、また、今後、電子的に情報収集を行う場合やビッグデータを利用する場合にも、同様の議論が生じる可能性があるのではないか、との報告がなされた。

第4報告者の林氏からは、（1）2015年9月に持続可能な開発目標（SDGs）が採択されたが、持続性そのものを測る指標が必要である。（2）現在、国連（UN Statistical）を中心にInter-agency and Expert Group on Sustainable Development Goal Indicators（IAEG-SDGs）が指標作りの作業中である。しかし、各分野別の従来から使われている指標を使用する方向で議論がなされており、持続性の問題に本格的に切り込んでいない。（3）持続性を指標化する場合には、国内総生産（GDP）のような付加価値ベースのフローの指標のみならず、枯渇性資源の使用などをストックの減少としてとらえて、指標化していく必要がある。グリーンNNPOは経済環境統合勘定などを導入する試みも長年行われているが、まだ本格的に世界共通の指標としては使われていない。（4）資源の収奪や環境の劣化を組み入れたマクロ指標を確立することがSDGs期間中の大きな課題である、との報告があった。

第5報告者の山谷氏からは、日本の政策評価（中央府省）、行政評価（地方自治体）の歴史について報告があった。1990年代に導入が検討され、制度化されて以来、この二つの評価システムには大きな問題が発生した。一つは公務員制度が原因の病であり、繰り返される評価担当者の人事異動が原因で発生する「経年劣化」である。もう一つは評価作業に予算編成作業を重ねた結果生じた「制度疲労」の病である。この二つによって国際的に見ればかなり異質の評価システムが現出した。このガラパゴス化の実際は、政策評価が目標管理型政策評価に変わり、行政評価は評価シートに記入する年中行事になった有様で現れた。もともと評価は政策やプログラムの修正･矯正、場合によっては事業廃止を招くため、終身雇用や年功序列型の日本公務員文化にとっては毒が多いシステムであったが、日本の公務員制度はこれを20年かけて「無毒化」した、との報告であった。

第6報告者の和田氏からは、ODA評価の実務的な課題の4点として、MDG支援の実施ツールであったPRSCからDPLなどの財政支援、政策借款の政策評価の手法とPDM/PCMへの導入の課題、国内研修・留学生（JDS）などの効果の評価の課題、インフラ輸出・質の高いインフラについての政策面からの評価の課題、一層進めば、安全保障の視点からのODA評価に関する課題があるとの報告があった。

これらの報告に対して会場からは、評価の独立性を完全に保つこと出来るのか、独立を保つことが実際的か、安全保障の視点からのODA評価はどうするか、などを含む多数の質問・意見が出された。日本における評価のあり方については、今後とも継続的に議論がなされる必要があろう。

企画委員会

企画委員会は、日本評価学会が行う各種の活動を企画、実行することをとおして、わが国の評価学の普及と発展に寄与することを目的としています。最近の活動としては以下の活動があります。

1. 評価士養成・認定

公共事業、学校教育、福祉、ODA等をはじめとする様々な分野の政策や事業に関するアカウンタビリティや効率化が求められており、こうした社会一般からの需要に応えるために、評価システムの導入や事業評価の実施が進められていますが、評価の専門性を持つ人材は未だ十分に育っていないのが現状です。そこで、評価に関する専門的能力を身につけた人材を養成してその能力を認定し、各分野における評価の向上に資することが目的です。

評価士はそのレベルによって初級、中級、上級に区分しています。それぞれのレベルの考え方は次のとおりです。

初級レベル（評価士）：　評価に関する幅広い知識と技法を身につけた専門家。

中級レベル（専門評価士）：各専門分野（セクター）の評価に関する深い知識と技法を身につけた専門家。

上級レベル（上級評価士）：評価に関する体系的な研究を行うことにより、評価活動の発展に寄与することができる専門家。

平成27年度後半～ 28年度前半には、初級レベルの第18、19期評価士養成講座を開講し、計60人が受講されました。また、平成28年5月に開催された春季第13回全国大会において、中級レベルの専門評価士に関して「「行政評価士」の役割とそのカリキュラム」と題して、議論が行われました。

2. 顕彰事業

日本評価学会では年一回、学会賞の選定という形で会員の顕彰を行うことにしています。学会賞の受賞者は企画委員会の下に設けられている学会賞審査委員会によって推薦し、理事会で決定されます。学会賞としては、本学会誌に掲載された原著論文等が評価学研究に大きく貢献したと認められる者に授与される「日本評価学会論文賞」、評価学研究の進歩に寄与する優れた研究をなし、将来の発展を期待される者に授与する「日本評価学会奨励賞」、評価学の発展に顕著な功績があった者に授与される「日本評価学会功績賞」の3つの賞があります。

平成28年度は10月28日まで受賞者の推薦を受け付けています。詳細は学会ホーム・ページに記載してあります。多くの自薦、他薦を期待しています。受賞者は所定の手続きを経て決定され、11月26、27日に広島大学で開催予定の第17回全国大会の際に表彰されます。

企画委員会委員長　牟田　博光

国際交流委員会

国際交流委員会では海外の評価専門家や評価学会との交流や意見交換を通じて相互に学習を行っている。APEA（Asia-Pacific Evaluation Association）の活動も活発になっており、APEAとの連携も深まっている。2015年10月から2016年6月までに以下の活動を行った。

1. 第16回全国大会で国際セッションを開催

2015年12月12、13日にJICA沖縄国際センターで開催された全国大会で「評価の独立性、有効性のためのガバナンス」をテーマに国際セッションを開催した。多角的な視点から有益な議論が展開した。

発表：　廣野 良吉（成蹊大学）
　　Naidoo, Indran A.（UNDP）
　　Rasappan, Arunaselam（APEA, Center for Development & Research in Evaluation, International）
コメント：佐々木 亮（国際開発センター）
　　上野 宏（国際開発センター）
司会：　湊 直信（国際大学）

2. 第13回春季全国大会で国際セッションを開催

2016年5月28日にJICA横浜国際センターで開催された春季全国大会で「政策決定のための系統レビューにおける国際的潮流」をテーマに国際セッションを開催した。特にEvidence Based Policymakingに焦点を当て、Skype参加のAPEAメンバーと共に多様な視点から議論を行った。

発表：　Gough, David（EPPI-Centre, University College London）
　　正木朋也（JICA）
コメント：米原あき（東洋大学）
　　佐分利応貴（総務省）
　　APEA参加者（skypeによる）

3.APEAの理事会開催

Skypeを使用した理事会が毎月開催されているが、湊APEA副会長が全ての理事会に参加した。

4. APEA conference

APEAでは初めての国際コンフェレンスを11月21 ～ 25日にベトナムのハノイで開催することを計画している。そのための準備（発表公募、応募の選考、共催・スポンサー集め、ホテル予約等）をAPEAの一員として実施している。

5. ヨーロッパ評価学会誌Evaluationへの寄稿

Issue 22.4（July 2016）のNews from the Communityコラムに、上記2の国際セッション開催を報告した。

国際交流委員会委員長　湊　直信

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。
その効果を高めるため、文字情報とともに、写真など視覚情報も充実していきます。
また、自治体、企業などの説明ツールとしては、マス・メディアを使ったパブリシティも利用する方向で検討を進めていきます。

(2)「広報の手段は主として学会ホームページを活用します。従って、学会からのおしらせはホームページに随時掲載します」。
引き続き、より読みやすく、タイムリーに情報をお伝えするよう検討していきます。

3. 今後の方針

広報委員会の役割について見直し、進捗状況管理、新企画の検討などを行い広報面での有効性、効率性を高めて行く所存です。

広報委員会委員長　高千穂　安長

日本評価学会第17回全国大会のご案内

第 17 回全国大会
実行委員長　石田　洋子

来たる 2016 年 11 月 26 日（土）、27 日（日）に、日本評価学会第 17 回全国大会が広島大学（広島県東広島市）にて開催されます。

プログラム等詳細につきましては、学会メーリングリストおよびホームページにて随時ご案内いたします。

会員各位におかれましては、奮ってご参加くださいますようよろしくお願い申し上げます。

記

１．　日程：2016年11月26日（土）－27日（日）

２．　場所：広島大学
〒 739-8511　広島県東広島市鏡山 1-3-2

以上

日本評価学会誌刊行規定

2005.2.15改訂
2002.9.18改訂
2001.9.9改訂

（目的および名称）

1. 日本評価学会（以下、「学会」という）は、評価に関する研究および実践的活動の成果を国内外の学界をはじめ評価に関心をもつ個人および機関に広く公表し、評価慣行の向上と普及に資することを目的として、「日本評価研究（仮名）」（英文仮名："The Japanese Journal of Evaluation Studies"、以下、「評価研究」という）を刊行する。

（編集委員会）

2. 「評価研究」の編集は、後で定める「編集方針」にもとづいて編集委員会が行う。
3. 編集委員会は、学会会員20名以内をもって構成し、委員は学会理事会が選任する。編集委員の任期は2年とし、再任を妨げないものとする。
4. 編集委員会は、互選により委員長1名、副委員長2名および常任編集委員若干名を選出する。
5. 編集委員会は、最低年1回編集委員会を開き、編集方針、編集委員会企画、その他について協議するものとする。
6. 編集委員会は、その活動等について、随時理事会へ報告し、承認を受けるとともに、毎年1回学会年次大会の場で、過去1年の活動成果と翌年の活動計画に関する報告を行う。
7. 委員長、副委員長および常任編集委員は、常任編集委員会を構成し、常時、編集実務に当たる。

（編集方針）

8. 「評価研究」は、原則として、年2回刊行する。
9. 「評価研究」の体裁は、B5版とし、和文又は英文とする。
10. 「評価研究」に掲載する原稿（以下「論文等」という）の分類は、以下の5カテゴリーからなるものとする。
    (1) 総説
    (2) 研究論文
    (3) 研究ノート
    (4) 実践・調査報告
    (5) その他
11. 「評価研究」への投稿有資格者は、学会会員および常任編集委員会が投稿を依頼した者とする。学会会員による連名での投稿および学会会員を主筆者とする非会員との連名での投稿は、これを認める。編集委員による投稿はこれを認める。
12. 投稿原稿を上記分類のどのカテゴリーとして扱うかは、投稿者の申請等をもとに常任編集委員会が、下記の「作業指針」に従って決定する。
    (1)「総説」は、評価の理論あるいは慣行について概観する論文とし、その掲載については編集委員会が企画・決定する。
    (2)「研究論文」は、評価の理論構築あるいは慣行の理解について重要な学問的貢献となると認められる論文とし、その採否については次項に定める査読プロセスを経て常任編集委員会が決定する。
    (3)「研究ノート」は、「研究論文」作成過程での理論的あるいは経験的な研究の中間的成果物に相当する論考で、その採否については次項に定める査読プロセスを経て常任編集委員会が決定する。
    (4)「実践・調査報告」は、評価事業の実践あるいは評価にかかわる調査の報告で、その採否については次項に定める査読プロセスを経て常任編集委員会が決定する。

(5)「その他」には、編集委員会が独自に企画する特集に掲載する依頼原稿や学会誌の刊行に関する編集委員会からの学会会員への連絡等が含まれる。

13. 論文等は2名の査読者により査読することとし、その人選は編集委員会が行う。「研究論文」については、査読結果と編集委員会が査読者とは別に指名する担当編集委員1名の参考意 見をもとに、編集委員会が掲載に関する決定を行う。「総説」、「研究ノート」、「実践・調査報告」および「その他」の論文については、査読結果にもとづき編集委員会が掲載に関する決定を行う。

14. 編集委員が「評価研究」に投稿した場合には、当該委員はその投稿に係わる常任編集委員会あるいは編集委員会の議事に一切参加しないものとする。

15. 上記いずれのカテゴリーの投稿についても、常任編集委員会による掲載の判断は可・不可の二者択一で行うこととする。但し、場合によっては編集委員会の判断で、小規模の修正による掲載も認める。「研究論文」としての掲載が適当でないと判断された場合でも、投稿者が希望すれば、常任編集委員会は「研究ノート」あるいは「実践・調査報告」としての掲載を決定できる。

(投稿要領の作成公表)

16. 編集委員会は、上記の編集方針にもとづき投稿要領を作成し、理事会の承認を得て、広く公表する。

(配布先)

17. 「評価研究」は、学会会員に無償で配布するほか、非会員に有償で提供する。

(抜刷の配付)

18. 「評価研究」掲載論文等の抜刷り30部を、投稿者（原著者）に無料で配布する。それ以上の部数を希望する場合は投稿者（原著者）の自己負担とする。

(インターネット上の公開)

19. 「評価研究」掲載論文等は、投稿者（原著者）の了承を得て全文をインターネット上で公開する。

(著作権)

20. 「評価研究」に掲載された論文等の著作権は各投稿者（原著者）に帰属するものとし、編集権は本学会に帰属するものとする。

(事務局)

21. 「評価研究」編集及び配布の事務は、それに関連する会計も含めて学会事務局が担当する。

（以上）

『日本評価研究』投稿規定

2008.9.29改訂
2003.4.18改訂
2002.3.25改訂
2001.9.9改訂

1. 『日本評価研究』（The Japanese Journal of Evaluation Studies）は、評価に関する論文、論考、調査報告等を掲載する。

2. 『日本評価研究』は、会員間の研究成果交流の場を提供し、内外における評価研究の一層の発展に資することを主目的として発行されており、原則として会員による寄稿を掲載する。なお、依頼原稿を除き、ファーストオーサーは学会員でなければならない。また、投稿は、一時に一原稿に限るとともに、他学会誌などへ二重に投稿などのない未発表のものとする。

3. 投稿された原稿は、編集委員会の責任において審査を行ない、採否を決定する。審査にあたっては、1原稿毎に2名の査読者を選定し、査読結果を参考にする。（査読者には、投稿者名を伏せて査読を依頼する。）

4. 原稿料は支払わない。

5. 『日本評価研究』に掲載された論文等は、その全文をインターネット上の本学会のホームページに掲載する。

6. 投稿にあたっては、投稿原稿が、①研究論文、②総説、③研究ノート、④実践・調査報告、⑤その他のうち、どのカテゴリーに入るかを明記する。ただし、カテゴリーについての最終判断は、編集委員会で行なう。「研究論文」は評価の理論構築あるいは慣行の理解について重要な学問的貢献となると認められる論文、「総説」は、評価の理論あるいは慣行について概観する論文、「研究ノート」は「研究論文」作成過程での理論的あるいは経験的な研究の中間的成果物に相当する論考、「実践・調査報告」は評価事業の実践あるいは評価にかかわる調査の報告、「その他」は編集委員会が独自に企画する特集に掲載する依頼原稿等である。

7. 投稿方法
   （1）使用言語は日本語又は英語とする。
   （2）著者校正は原則として第一校までとする。
   （3）英文原稿については、ネイティブスピーカーによる英文チェックを済ませ、完全な英文にして投稿すること。
   （4）ハードコピー 4部（A4版）を提出する。その際、連絡先（住所、Tel、Fax、Email）と原稿の種類を明記すること。掲載可と判断された原稿については、必要なリライトを経た後に、最終原稿のハードコピー 2部とDOS/Vフォーマットのフロッピーを用いたTEXTファイルを提出する。その際、オリジナル図表を添付すること。
   （5）刷り上がりは最大14ページとする。これを超える場合は、その経費は著者負担とする。

（6）日本語原稿の最大文字数は以下のとおり。①研究論文20,000字、②総説15,000字、③研究ノート15,000字、④実践・調査報告20,000字、⑤その他適宜。それぞれ和文要旨を400字程度、英文要旨を150words程度、及び和文・英文でキーワード（5つ以内）を別に添付する。印刷は1ページ、20字×43行×2段（1,720字）とする。20,000字の原稿の場合、単純計算では英文要旨1ページを加えて合計13ページとなるが、図表の量によっては、それ以上のページ数となり得るので、注意すること。

（7）英文ではA4版用紙に左右マージン30mmをとり、10ポイントフォントを使用し、1ページ43行のレイアウトとする（1ページ約500words）。論文冒頭に150words程度のAbstractをつける。14ページでは、7,000words相当になるが、タイトルヘッド等を考慮して、最大語数を約6,000words（図表、注、文献込み）とする。図表の量によっては、ページ数が予想以上に増える場合もあり得るので、注意すること。

8.　送付先

〒108-0075　東京都港区港南1-6-41　品川クリスタルスクエア12階
一般財団法人 国際開発センター内 日本評価学会事務局
TEL: 03-6718-5931
FAX: 03-6718-1651
E-mail: jes.info@evaluationjp.org

『日本評価研究』執筆要領

2002.9.18改訂
2002.3.25改訂

1. 本文、図表、注記、参考文献等

（1）論文等の記載は次の順序とする。

日本語原稿の場合

第1ページ：表題、著者名、所属先、E-mail、和文要約（400字程度）、和文キーワード（5つ以内）
第2ページ以下：本文、謝辞あるいは付記、注記、参考文献
最終ページ：英文表題、英文著者名、英文所属先、E-mail、英文要約（150words程度）、英文キーワード（5つ以内）

英文原稿の場合

第1ページ：Title; the author’s name; Affiliation; E-mail address; Abstract（150 words）; Keywords（5 words)
第2ページ以下：The main text; acknowledgement; notes; references

（2）本文の区分は以下のようにする。

例１（日本語）

1.
（1）
①
（2）
（3）

例２（英文）

1.
1.1
1.1.1
1.1.2

（3）図表については、出所を明確にする。図表は原則として、筆者提出のものをそのまま写真製版するので、原図を明確に作成すること。写真は図として扱う。

例１：日本語原稿の場合

図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）おとしで記述する。

『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。

Publication Policy of the Japanese Journal of Evaluation Studies

Last revised on 15th February 2005

The Purpose and the Name

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

Editorial Board

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

Editorial Policy

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

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

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

12.1. "Review" is a paper, which provides an overview of evaluation theory or practice. The editorial board will make the decision regarding publication.

12.2. "Article" is considered as a significant academic contribution to the theoretical development of evaluation or understanding of evaluation practice. The standing editors committee makes adoption judgments following the referee-reading process described in the next section.

12.3. "Research note" is a discussion equivalent to the intermediate outputs of a theoretical or empirical enquiry. The standing editors committee makes adoption judgments following the referee-reading process described in the next section.

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

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

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

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

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

Formulation and Release of Submission Procedure

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

Distribution

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

Distribution of the Printed Manuscript

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

Release on the Internet

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

Copyright

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

Office

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

Information for Contributors
(For English Papers)

Last revised on 29th September 2008

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

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

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

4. Payment for the manuscript is not provided.

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

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

7. Manuscript Submission
   (1) Manuscripts may be written in either Japanese or English.
   (2) Correction by the author is only for the first correction.
   (3) English manuscripts should be submitted only after the English has been checked by a native speaker.
   (4) Submit four hard copies (A4 size) of the manuscript. Contact information including mailing address, telephone number, fax number, and e-mail address, and the category of the manuscript should be clearly stated.
   For approved manuscripts, after necessary rewriting, the author needs to submit two hard copies of the final paper as well as a text file saved on a DOS/V formatted floppy disk. Original figures, charts, and maps should be provided.
   (5) Total printed pages should not exceed 14 pages. Any cost incurred by printing more than 14 pages must be covered by the author.
   (6) The layout for English papers should be 30 mm of margin at left and right side, 10pt for font size, 43

lines on A4 paper (about 500 words per page). An abstract of 150 words should be attached to the front. 14 pages are equivalent to 7,000 words but the body should not exceed 6,000 words to allow for the title, header, figure, chart, footnotes, and references. Please note that the number of pages may be more than expected depending on the number of figures included.

8. Mailing address

Office of Japan Evaluation Society at International Development Center of Japan
Shinagawa Crystal Square 12th Floor, 1-6-41 Konan, Minato-ku, Tokyo,
108-0075, Japan
Phone: +81-3-6718-5931, Facsimile: +81-3-6718-1651
E-mail: jes.info@evaluationjp.org

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

Revised on 18th September 2002

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

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

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

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

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

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

Figure 1 Number of Students in the State of ○○

Note:
Source:

Table 1 Number of Accidents in the State of ○○

Note:
Source:

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Referee-Reading Guideline

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

1. Content of the Referee-Reading Guideline

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

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

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

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

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

3. Items of Consideration in Referee-Reading

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

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

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

4. Attentions in submission of manuscript

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

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

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

5. Judgment Cases in Referee-Reading

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

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

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

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

6. Judgment

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

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

～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。随時、投稿を受け付けておりますので、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　投稿窓口
〒108-0075　東京都港区港南1-6-41品川クリスタルスクエア12階
一般財団法人国際開発センター内
e-mail : jes.info@evaluationjp.org

『日本評価研究』第16巻第1号
2016年8月31日

編集・発行　特定非営利活動法人　日本評価学会
〒108-0075　東京都港区港南1-6-41
品川クリスタルスクエア12階
一般財団法人国際開発センター内

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。

ISSN 1346-6151

Japanese Journal of Evaluation Studies

Vol. 16, No. 1, August 2016

CONTENTS

Special Issue: Two Decades of Evaluation Practices in the Japanese Local Governments

Editor's Note: Local governance reforms and evaluations in Japan — Kiyoshi Yamaya

Current State and Future Agenda for Performance Measurement in the Local Government: Lessons Learned from 20 years Practices in the 47 Prefectures — Tatsuya Ono

Looking back on Performance Evaluation System in Local Government: Two Decades of the Non-utilization nor Discontinuance — Junro Nishide

Many faces of evaluation in Japanese local governments -20 years of Decentralization and the Financial crisis- — Kiyoshi Yamaya

Research notes

Networked Public Services: Its Management, Governance, and Performance Measurement -Reviews of the Literature and Agenda for the Future- — Jun Arakawa, Masatoshi Tamamura

Report of the 12th Spring Conference

Report of the 16th Annual Conference

Announcement for the 17th Annual Conference

Report of the Committees

Japan Evaluation Society