

日本評価研究

Japanese Journal of Evaluation Studies

Vol. 13, No. 2, November 2013

特集：政策評価制度10年の経験～レビューと展望～

特集：「政策評価制度10年の経験～レビューと展望～」 山谷 清志

政策評価制度10年の軌跡
～制度導入以降の省察と今後の展望～ 新井 誠一

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～ 小野 達也

政策評価制度における評価の利用：
「評価の利用」概念の解明と現状の分析 田中 啓

政策評価とアカウンタビリティ
～法施行後10年の経験から～ 南島 和久

実践・調査報告
独立行政法人の二次評価に関する考察 竹中 一人

『日本評価研究』編集委員会
Editorial Board

編集委員長
Editor-in-chief

三好 皓一(立命館アジア太平洋大学)
Koichi MIYOSHI

副委員長
Vice-Editor-in-chief

西野 桂子(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

山谷 清志(同志社大学)
Kiyoshi YAMAYA

編集委員
Editors

青山 温子(名古屋大学)
Atsuko AOYAMA

岩渕 公二(NPO政策21)
Koji IWABUCHI

大島 巖(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

佐藤真理子(筑波大学)
Mariko SATO

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

〒140-0002 東京都品川区東品川4-12-6 日立ソリューションズタワーB 22F
一般財団法人国際開発センター内
特定非営利活動法人日本評価学会
編集事務担当 鈴木 直美
TEL: 03-6718-5931, FAX: 03-6718-1651
E-mail: jes.info@evaluationjp.org

『日本評価研究』は、和文年2回、英文年1回または2回刊行します。

日本評価研究

第13巻 第2号 2013年11月

目次

特集：「政策評価制度10年の経験～レビューと展望～」

山谷 清志

特集：「政策評価制度10年の経験～レビューと展望～」 1

新井 誠一

政策評価制度10年の軌跡

～制度導入以降の省察と今後の展望～ 3

小野 達也

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～ 21

田中 啓

政策評価制度における評価の利用：

「評価の利用」概念の解明と現状の分析 37

南島 和久

政策評価とアカウンタビリティ

～法施行後10年の経験から～ 53

実践・調査報告

竹中 一人

独立行政法人の二次評価に関する考察 69

日本評価研究刊行規定 83

日本評価研究投稿規定 85

日本評価研究執筆要領 87

日本評価研究査読要領 90

Publication Policy of the Japanese Journal of Evaluation Studies 92

Information for Contributors (For English Papers) 94

Writing Manual of the Japanese Journal of Evaluation Studies (For English Papers) 96

Referee-Reading Guideline 98

【巻頭言】

特集：政策評価制度10年の経験 ～レビューと展望～

山谷 清志

同志社大学

日本の政策評価制度は1997年12月、行革会議『最終報告』において「評価制度の導入」が提言されたことに始まる。その後、1998年の中央省庁等改革基本法において政策評価機能強化が謳われ、1999年には内閣府設置法、国家行政組織法改正、総務省設置法によってさらに念を押す形で政策評価の推進が求められた。それらを受けて2001年に各府省が組織令によって政策評価担当組織を設置し、最後に2001年6月「行政機関が行う政策の評価に関する法律」が成立して、この制度は一応の完成を見る。

ただし政策評価制度は当初、各府省にその運用が任されていたこともあって、採用される評価方式、外部有識者会議の運用方法、評価シートや評価票の形式や記述内容において、かなり多様なスタイルがとられていた。もちろん評価書を公表するに際しても、冊子形態で刊行する省もあれば、ホームページでの公表にとどまるところもあった。また当初評価件数が1万件を超え、さまざまな政策評価が膨大な量、毎年公表されることに戦慄を覚えることすらあった（後に重点化され件数は大幅に削減された）。

この政策評価制度を一定の形に集約させようとする努力は、2003年の「経済財政運営と構造改革に関する基本方針2003」（2003年6月27日閣議決定）から始まる。この基本方針では、次の三点が主張された。

- 定量的な達成目標、達成期限、達成手段が明示されていること。何をもって「達成」とするか、評価方法が提示されていること。目標期間は1～3年程度とし、各年度ごとの達成目標が明らかにされていること。
- 政策目標を効率的に達成するため、事業の性格に応じ、予算執行の弾力化を行い、各府省は、弾力化に伴う効率化に応じ、これを予算に反映する。
- 複数年度にわたる予算執行に支障のないようにする。

それから10年、定量的な目標の設定と予算への反映が内閣の重要なミッションになり、そのためにさまざまな調査研究にもとづいた制度の修正が行われた。早くは2004年1月、内閣府の大田弘子・大臣官房審議官（当時）はスウェーデン、オーストラリア、ニュージーランドで調査を行い、また同じく寺田稔・内閣府参事官（当時）もイギリス、ドイツ、フランスでヒアリングを行った。その結果は2004年度の「モデル事業」に反映された。すなわち、明確な数値目標を立てて複数年度にわたって行う事業については弾力的に予算を執行する、執行のさまざまな手続は簡略化するが事後評価は的確に実施する、このようなアイデアをモデル事業として導入したのである。こうした方向は政権交代後の民主党政権に代わってもさらに明確に打ち出され、菅直人副総理は「目標管理型の政策評価」として指示（2009年9月）、各府省の実務担当者には「目標管理型の政策評価にかかる評価書の標準様式」にもとづく評価と、事前分析表の作成が求められることになる。政治家が熱心なマネジメント・サイクルでの活用方針が、いつの間にか評価をマネジメント・ツールと言い換えさせ、評価と測定が同義になっていった。

こうして、日本の政策評価はかなりユニークな姿に集約されようとしている。すなわち、定量目標の設定、その達成度の測定、測定結果を予算編成に反映するマネジメント型モニターの姿である。しかしそれが本来の政策評価であると言うのは難しい。評価の理論で言えば、この方法は限りなくパフォーマンス測

定（performance measurement）に近く、測定になじむ対象は政策ではなく事業であり、また成果主義を標榜しながら実際には活動指標に落ち着くことが多い。行政サービスを提供する組織のトップダウン管理のイメージが強いので、政策評価の外部有識者会議での議論は精緻な業績指標の設定や予算への反映に集中する。その上、政策評価とは似て非なる行政事業レビューを意識せざるを得ない制度運用も始まっている。評価研究者や実務担当者が苦勞するこの状況は、実は日本だけではなく長い歴史を持つアメリカにおいても同じように発生していた。すなわちアメリカ評価学会の機関誌のひとつ*New Directions for Evaluation*の2013年春号においても、専門が違い、知的背景も違い、文化も違う業績測定と評価との棲み分け、あるいは相互補完関係のあり方を考える特集‘Performance Management and Evaluation’が組まれ、業績測定をマネジメントに使う場合、評価をどのように考えるべきなのかが議論されている。

この特集「政策評価制度10年の経験～レビューと展望～」もまた、意図しなかったにもかかわらずアメリカ評価学会の特集と同じ視点を持っており、実務担当者と評価の研究者が以下のような議論を展開する。

新井誠一「政策評価制度10年の軌跡～制度導入以降の省察と今後の展望～」

小野達也「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～」

田中 啓「政策評価制度における評価の利用：『評価の利用』概念の解明と現状の分析」

南島和久「政策評価とアカウンタビリティ～法施行後10年の経験から～」

政策評価をどのように使うのか、また独立行政法人評価をはじめとする他の政府内の評価とどのように使い分けるのか、再度問われている現状の中で、日々悩んでおられる実務家の方がたには申し訳ないが、興味は尽きない。

【研究論文】

政策評価制度10年の軌跡 ～制度導入以降の省察と今後の展望～

新井 誠一

総務省行政評価局

sarai@soumu.go.jp

要 約

21世紀の到来とともに、平成13年1月、我が国の行政に政策評価制度が導入されて約12年が経過した。

政策評価制度は、プラン偏重の行政を改善すべく、総務省に設置された政策評価の手法等に関する研究会が検討を進め、標準的ガイドラインを策定してスタートした。

その後、制度は法制化され、各府省のマネジメントサイクル（PDCAサイクル）の中で定着したが、法施行から3年後の平成17年の時点で制度全般にわたる見直しが行われ、平成21年に事業仕分けを経て抜本的機能強化を図ることとされ、目標管理型の評価の導入など現在に至っている。

一方で、依然として政策評価の現場などで、評価方式、評価結果の活用などについて様々な論点が指摘されている。

本稿では、制度導入以来様々な見直しが行われたにもかかわらず、こうした指摘がなされている原因を政策評価制度の導入から制度見直しを経て現在に至る経緯の中に探り、制度の今後の展望を明らかにしたい。

キーワード

政策評価制度、マネジメントサイクル、制度全般にわたる見直し、抜本的機能強化

はじめに

21世紀の到来とともに、平成13年1月、日本の行政に政策評価制度が導入されて、約12年が経過した。また、平成14年4月には制度の実効性・信頼性を確かなものにするために「行政機関が行う政策の評価に関する法律」（以下、「評価法」という。）が制定された。

平成17年4月には、評価法の施行から3年が経過したことを受け、法に置かれた3年経過時の「見直し条項」に基づき、政策評価制度の全般にわた

る見直しが行われた。その後、平成21年に民主党を中心とした政権の下、政策評価制度も行政刷新会議による事業仕分けの対象となり、その結果抜本的機能強化を図ることとされた。

さらに、現在の自公政権の下では、経済財政諮問会議や行政改革推進会議で政府内のPDCAサイクル（本稿において政府のマネジメントサイクルを言う。）の強化という観点から政策評価が取り上げられている。

① 一方で、制度導入から10年余りが経過し、様々な見直しを経た現時点でも政策評価の現場

や経済財政諮問会議等の政府の会議の場で、以下のような政策評価制度をめぐる論点が指摘されている。評価方式について、様々なものが用いられているが、制度を論ずるときにどの評価方式を取り上げているか整理せずに、特定の方式を前提として全体的な批判を行う場合がある。

- ② 評価結果の活用について、政策評価は、あくまで政策の改善という意味決定のための判断材料を提供するのが本来の役割であるが、評価結果をもって自動的に政策判断ができるとの過大な期待を持って議論される場合がある。
- ③ 評価目的について、評価結果は、政策の廃止・縮小に結びつく場合もあれば、充実につながる場合もある。予算の削減につながらなければ意味がないとの議論がなされる場合がある。
- ④ 評価対象について、特に制度導入当初はできる限りカバーすることを重視し、かえって総花的な評価になった面がある。
- ⑤ 評価時期について、政策の企画立案に際しては事前評価が必要であること、政策の実施の最中にもプロセスチェックがあることなど、これらの過程は実際には渾然一体となって進んでいるが、政策の実施が終わってから評価が始まるとの考え方が一般的である。

本稿では、制度導入後様々な見直しを経たにもかかわらず、こうした論点がなぜ現在も指摘されるに至っているか、その原因を政策評価制度の導入から制度見直しの経緯の中に探り、政策評価制度の今後の展望を明らかにしたい。なお、本稿中、意見にわたる部分は、筆者の私見であることを予めお断りする。

1. 政策評価制度導入の背景

政策評価は我が国のみならず、多くの国々において高い関心の下で推進されている取組である。1970年代における2度の石油危機等を契機として、主要OECD諸国においては経済成長が停滞し、税収減に伴う累積財政赤字の増大が共通する課題として認識されるようになった。¹各国はこれらの財政赤字の解消のため公共支出の削減と、それに伴う公共部門の在り方の見直しを迫られた。そこ

では、限られた資源のもとでの効率的・効果的な公共サービスの供給が求められていた。

我が国においては、平成9年の行政改革会議の最終報告において全政府的な政策評価制度の導入が盛り込まれたことが契機となって、制度の導入が進められることとなったが、制度構築に当たっては、米国、英国、カナダ、オーストラリアなどで現地調査を行うなど、諸外国の先進的な取組を参考にしたところである。

以下、我が国における政策評価制度導入までの経緯を具体的に振り返る。

(制度の経緯は図1参照)

2. 政策評価制度導入の経緯

(1) 行政改革会議最終報告から中央省庁等改革基本法へ

① 行政改革会議からの問題提起

21世紀の到来を控え、政府では中央省庁再編が課題とされ、平成8年、内閣総理大臣をトップとする行政改革会議が設置された。同会議の日本の行政に対する問題意識として、従来、我が国の行政において法律の制定や予算の獲得に重点が置かれ、その効果やその後の社会経済情勢の変化に基づき政策を見直すといった評価機能は軽視されがちであった、との意見が示された。

当時は、諸外国でも行政活動に対する事後評価を重視する考え方が採り入れられつつあり、こうした動向も視野に入れて、議論が展開された。その結果、平成9年12月に最終報告がまとめられ、その中で、全政府的に政策評価制度を導入することが盛り込まれた。

② 中央省庁等改革における具体化作業

行政改革会議の最終報告を具体化するための組織としては、中央省庁再編が全政府的な重要な事項であることから、内閣官房に中央省庁等改革推進本部が設置され、従来の行政監察機能を担ってきた実績からノウハウ・人材を擁する総務庁が同本部に協力することとされた。

こうした体制の下で政策評価制度の導入に向けた具体化作業として法的整備に着手し、まず、平成10年6月に成立した「中央省庁等改革基本

法」の中で「国民的視点に立ち、かつ、内外の社会経済情勢の変化を踏まえた客観的な政策評価機能を強化するとともに、評価の結果は政策に適切に反映されるようにする」ことが規定された。また、各府省における政策評価機能の充実強化を図るための措置として、①明確な位置付けを与えられた評価部門の確立、②府省の枠を超えた政策評価機能の強化、③政策評価に関する情報の公開、④評価結果の政策への反映についての説明責任の明確化、などが規定された。さらに、総務省の編成方針として、行政の評価及び監視の機能について規定された。

これを受け、国家行政組織法等が改正され、各府省の基本的作用として、それまでの政策の企画立案（plan）・実施（do）に加えて、自ら評価（see）を行うことが明記された。このことは、評価の位置付けを明確にする上で重要な措置であり、また、いわゆる「plan-do-see」のマネジメントサイクル（本稿でのマネジメントサイクルは、「PDCAサイクル」と「PDSサイクル」の両方を言う。）が我が国法制上、確立されたことを意味する画期的な法的措置であった。なお、この点は、本稿「はじめに」で掲げた論点⑤との関連で押さえておく必要がある。

なお、これらの法案の国会審議の結果、中央省庁等改革基本法の附帯決議において、「行政評価法（仮称）制定について検討」することが盛り込まれた。この点について考察すると、国家行政組織法等の改正により、各府省の基本的作用として評価を行うことが盛り込まれたことで、政策評価制度は、特段の個別法を設けずに、政府内のガイドラインに相当する決定がなされれば法制度的には可能であり、実際、当時の政府はその前提で導入準備を進めていた。しかし、政策評価制度が、国会の行政に対するチェック機能という側面を有するものであることから、国会としては、このチェック機能を十全なものとするためにも、改めて特段の作用法を制定することを政府に求めたものであると考えられる。²

③ 総務省の役割

政策評価制度については、各府省の枠を超えて政策評価の総合性や一層厳格な客観性を担保

する役割を担う機関が必要とされ、これに適する機関として、それまでの行政監察機能を担い、ノウハウと人材を擁する総務省がふさわしいと判断され、このための規定が総務省設置法に位置付けられた。この結果、省庁再編後、総務庁行政監察局は、総務省行政評価局に改組され、行政監察、行政相談というそれまで担ってきた機能を継承するとともに、政策評価、独立行政法人評価という省庁再編を経て新たに担うこととされた機能が付加されて強化されることとなった。

行政評価局では、3年間で実施する予定の政策の評価や行政評価・監視のテーマを行政評価等プログラムに盛り込み、計画的に推進することとしており、毎年度ローリング方式で改定しているが、その中で統一性又は総合性を確保するための評価を実施することとしている。

(2) 研究会設置からガイドライン策定まで

① 研究会の設置

中央省庁改革基本法を具体化するため、平成11年4月に中央省庁等改革推進本部で決定された「中央省庁等改革の推進に関する方針」の中で、省庁再編後に政策評価制度を所管することが既に決まっていた総務省において、政策評価の実施要領、評価基準等の標準的ガイドラインを作成することが新たに決定された。これを受け、当時の総務庁行政監察局に「政策評価等推進準備室」が設置され、標準的ガイドラインの策定に向けた検討が開始された。準備室は、当時の総務庁のスタッフ以外に、各省庁からの出向者で構成され、各省庁の業務の実態を把握しつつ、全政府的な協力を確保する方針で編成され、発足した。

しかし、当時は、公共事業、研究開発、政府開発援助（ODA）といった個別分野を除き、国レベルでの評価制度は存在しておらず、その設計を検討するに当たっては、いわば「ゼロベース」からスタートするに等しいという状況であった。このため、制度の検討に当たっては、政府内にとどまらず、広く知見を外部の有識者に求めることとし、学界、公認会計士、シンクタンクから7人の有識者の参集を求め、村松岐

夫京都大学大学院法学研究科教授（当時）を座長として、平成11年8月に「政策評価の手法等に関する研究会」を立ち上げ、その運営を通じて検討を深めるとともに、これをベースとして政府としてのガイドラインを策定するという形で進められた。

その際、各省庁の先行的な取組や今後の取組方針を調査やヒアリングなどで把握することはもとより、既に評価の取組が先行している諸外国や地方公共団体における取組事例を参考にしつつ、国のシステムとして導入するに当たって参考となる要素を幅広く取り入れ、また、適切な要素を組み合わせることで体系化するというアプローチが採られた。³その結果、徐々にではあるが、評価システムの骨格や評価の具体的な手法が明らかになっていったのである。筆者も事務局に携わったが、「まずは内外の知恵を総結集する」との姿勢で臨んだことが、ゼロベースからの検討に活路を見出したのである。

このことは、日本版政策評価制度が、政策全般を対象とするパッケージである面、評価対象に応じて適材適所な方式を組み合わせるコンビネーションである面といったユニークな特徴を有するに至った原因にもなった。なお、この点は、本稿「はじめに」で掲げた論点①との関連で押さえておく必要がある。また、当時、政策評価制度が行政改革の柱としての位置付けを有していたことを勧案すると、特定分野に限定されることなく、できる限り広範な政策を対象として評価を行っていくというこのアプローチを採ることが求められたとも考えられる。ただし、その後の制度の運用実態を見ると、あまりにも網羅的に評価が行われることで、真に評価の実施が必要とされる重要な政策について、十分な内容の評価がなされていないとの指摘があり⁴、評価の重点化・効率化は、政策評価制度の見直しの際の主要論点の一つとなった。なお、この点は、本稿「はじめに」で掲げた論点④との関連で押さえておく必要がある。

② 研究会最終報告で評価方式を提示

「政策評価の手法等に関する研究会」は、メンバーがそれぞれ項目を分担して報告を行う中で明らかになった論点を整理するとともに、各

方面からのヒアリングを精力的にこなし、「中間整理」、「中間まとめ」を経て、平成12年12月に「政策評価制度の在り方に関する最終報告」を発表した。その内容は、極めて包括的で実効性のある内容となった。正に、その後の政策評価制度の在り方の礎となったものであり、その基本的な精神は10年以上を経た今日の制度の中でもいまだに維持されており、いわば「政策評価の原点」と言っても過言ではないといえる。

その内容は、政策評価の基本的な在り方と政策評価の標準的な方式の導入及び実施の在り方について明らかにしている。さらに、政策評価の定着・発展に向けた課題の提示を行うなど、1年5か月にわたる21回の会合の集大成となっている。まず、政策評価の基本的な在り方として、導入の目的を、①国民に対する行政の説明責任（アカウンタビリティ）、②国民本位の効率的で質の高い行政の実現、③国民的視点に立った成果重視の行政への転換、としている。この目的は、いずれもあえて「国民」との関係に着目して打ちたてたものであり、政策評価制度は決して霞が関の中だけの制度ではなく、そこで行われている行政活動の対象となる国民との関わりが大きいことを強調している。

次に、政策評価の基本的な枠組みとして、①政策評価の概念、②政策評価の対象範囲、③政策評価の実施主体（各府省、総務省）、④政策評価の時点（事前、途中、事後）、⑤政策評価の観点等（必要性、効率性、有効性など）、について明らかにするとともに、政策評価の結果の反映、政策評価の結果等の公表、政策評価の方式、政策評価の手法について整理し、政府の標準的ガイドラインに盛り込むよう提言した。

このうち、政策評価の実施主体については、①対応すべき行政課題を最も把握しやすい立場にある各府省が自ら評価を行い、その結果を自ら企画立案やそれに基づく実施に反映させることで、実効ある改善・見直しが行われ、政策の質の向上が図られること、②各府省は、所掌する政策について最も詳しい情報・データを入手し得る立場にあること、③各府省が政策評価を通じて得られる知識・経験を学習・蓄積し、今後の政策の企画立案にいかしていくという過程

の確立によって、自らの政策形成能力が高まること、から各府省の自己評価を基本としている。一方で、いわゆる「お手盛り」評価にならないように評価の客観性を確保することや、各府省にまたがる政策についての評価を行うことが求められ、評価専担組織としての総務省が統一性・総合性確保評価を行うこととされた。

また、政策評価の方式については、対象となる政策のレベルや評価に求められる内容を踏まえて、「事業評価」、「実績評価」、「総合評価」の3方式を標準的な方式として我が国で初めてネーミングして打ち出した。他の項目についても、包括的・体系的に整理され、具体的な内容を盛り込んでおり、その後の政府の検討を牽引したと評価できる。なお、評価方式の選択は、各府省の自立的な判断に委ねられていたが、網羅的な実施を重視して実績評価方式を採用するところが結果的には多かった⁵。

なお、総務省が行う政策の評価について、現在の類型である①政策の統一性・総合性を確保する評価、②各府省が行った評価の客観性を担保するための評価、の大枠が示されるとともに、総務省に制度の基本的な事項や総務省が行う政策の評価に関する調査審議を行う第三者機関として、政策評価・独立行政法人評価委員会を置くことが盛り込まれた。

さらに、研究会は、政策評価制度が新たに導入されるものであることから、制度の定着・発展に向けた取組が必要であるとして、中長期的な宿題として、①職員の意識改革、②成果重視の徹底とインセンティブの付与、③人事管理との関係、④情報・データの整備、⑤予算との関係、⑥行政活動に伴うコストの把握、⑦研究開発の推進、教育研究の展開、⑧多元的な評価、を掲げている。⁶

③ 標準的ガイドラインの決定

政府は、「政策評価の手法等に関する研究会」の検討に併せて、並行的に標準的ガイドラインの策定作業を進めてきたが、研究会の最終報告をはばすべての項目にわたり反映した上で、正式に平成13年1月に標準的ガイドラインが政策評価各府省連絡会議で了承され、省庁再編に合わせて政策評価制度をスタートするという政府

の目標が達成された。

中央省庁新体制の下で、各府省には政策評価を専門に担う部局が官房に設置されるとともに、評価を行うに当たって知見を活用する有識者会議が設置されるなどの体制が整備された。また、同年3月までの年度内に評価を実施するための計画づくりが進められ、平成14年度から、順次、評価が開始された。また、総務省には、行政評価局が設置され、平成13年度内に制度の円滑・効果的な実施に向けた条件整備を行うとともに、自らが行う評価の実施に向けた計画を「行政評価等プログラム」として策定し、平成14年度から実際の評価活動に入った。

(3) 法制化へ

① 法制化の流れ

こうして政府における標準的ガイドラインの下での制度導入の作業は着々と進められたが、一方で、与党からは、政府の自己評価制度に対する実効性や国民からの信頼性を確保するには、法律という国会の意思決定を経る形式が必要であるとの意向が示された。こうした要請を受け、平成12年7月には、当時の森内閣総理大臣から、政策評価制度に関する早期の法制定について指示がなされ、法制化への流れが急浮上することとなった。⁷

これを受け、当時の総務庁では、標準的ガイドラインの策定時と同様に、外部の有識者の参集を得て、工藤敦夫元内閣法制局長官を座長として、同年9月に「政策評価制度の法制化に関する研究会」を立ち上げた。同研究会は、わずか2か月半程度で報告を取りまとめるという急ピッチな作業が行われた。これは、「政策評価の手法等に関する研究会の検討がゼロベースから開始されたのに対し、同研究会の検討の蓄積があったことも大きく寄与していると考えられる。

報告では、法制化に当たっては、標準的ガイドラインを基本としつつ、①政策評価の実施を法律上明確に義務付けるとともに、その着実な実施を図ること、②政策評価の客観性を確保すること、③政策評価に関する一連の情報の公表を義務付けること、に留意し、各府省の政策全般を対象とした政策評価制度の共通ルールを定

めるものと位置付けている。⁸

② 法案の提出

総務省は、この最終報告を踏まえ、「行政機関が行う政策の評価に関する法律案」を作成し、平成13年の通常国会に提出した。同法案は、衆・参両院の総務委員会での審議を経て、全会一致で可決・成立した。これを受け、事前評価が義務付けられる政策や事後評価が義務付けられる未着手・未了の期間などを定めた評価法の施行政令、また、各府省の基本計画の指針となるべき事項（観点、政策効果の把握、第三者の知見の活用、評価結果の政策への反映、評価関連情報の公表など）、評価の質の向上のための施策（職員の研修、調査研究、情報所在案内など）を内容とする基本方針が策定され、政策評価制度は、平成14年4月、評価法という新たなフレームの下でバージョンアップされた。なお、評価法の国会審議の過程で、評価制度が新しい制度であることから、附則第2条として、「法の施行から3年を経過した場合、法の施行状況を検討して、必要な措置を講ずる」とするいわゆる「見直し条項」が追加され、これに基づいて政策評価制度の見直しが行われることになった。

（法の骨格は図2参照）

3. 政策評価制度の見直し

（1）制度見直しの経緯

評価法の見直しに向け、まず、政策評価制度について審議を行う総務大臣の諮問機関である総務省の政策評価・独立行政法人評価委員会（以下、「委員会」と言う。）が始動した。具体的には、委員会の下に置かれた政策評価分科会において、見直し時期を1年後に控えた平成16年4月、丹羽宇一郎分科会長（伊藤忠商事（株）会長。その後、委員会の委員長に就任。）において、各府省と数回にわたり取組状況についての意見交換を行い、同年末に「政策評価制度に関する見直しの論点整理」を発表した。論点整理では、「効果的・効率的な行政の推進」、「政府の国民への説明責任の徹底」という評価法の目的を実現すべく、①評価の計画的・着実な推進、②客観的・厳格な評価の実施、

③評価結果の政策への反映、④評価に関する情報の公表、に必要な方策を検討することが必要であるとするとともに、総務省が担うべき役割について提示している。

① 総務省の検討

委員会の検討を受け、総務省では、政策評価制度の見直しの取りまとめを担う立場から、政策評価の実施状況と評価結果の政策への反映状況を把握するとともに、委員会と連携して対応方策の検討を進め、見直し時点からできる限り早期にその方向性を定めるべく検討を進め、政府内の調整を進め、「政策評価制度に関する見直しの方向性」を平成17年6月に発表した。その中では、今後の課題と対応方策として、①評価結果の予算要求等政策への反映、②重要政策に関する評価の徹底、③評価の客観性の確保、④国民への説明責任の徹底、⑤政策評価の基盤整備、を図ることとし、平成17年内にこれらの対応方策を盛り込んだ「政策評価に関する基本方針」の改定や、新ガイドラインの策定などを行い、平成18年度の評価から実施に移すこととした。

なお、委員会は、論点整理を発表した後も、この見直しの方向性を取りまとめるまで精力的に審議を行うなど、政府の制度見直し作業を牽引した。このような責任官庁とその第三者機関の連携プレーが功を奏したことは、見直し作業を円滑で実り多いものとした最大の要因であると考ええる。⁹

② 経済財政諮問会議の議論

経済財政諮問会議は、有識者議員を中心に、予算制度改革をテーマに取り上げてきたが、その一環として予算制度改革におけるモデル事業や政策群において政策評価の積極的活用を唱えるだけでなく予算・決算と政策評価の連携の強化の重要性を指摘してきた。

見直しの時点の半年前に当たる平成16年10月に有識者議員から、総務省に対して、①評価の様式を簡便で統一的なスタイルにすること、②達成目標の定量化の徹底、データの明示、③優先度を明らかにする仕組みの採用、④第三者評価の実施と結果の公表、を求めた。また、財務大臣に対しては、①政策評価調書の改善、②政

策と予算項目を結び付け把握できるような予算書の本格的見直しの着手、を求めた。これらの検討は、それぞれ年度内に中間的な取りまとめを行い、報告するよう宿題事項とされた。両省は、相互に連携して検討を行い、平成17年の3月10日、同時に検討状況を経済財政諮問会議において報告した。

これらの検討結果は、平成17年6月に閣議決定された「骨太方針 2005」において、①政策ごとに予算と決算を結びつけ、予算と成果を評価できるよう、予算書、決算書の見直しを行う、②政策評価と予算の連携強化を含め、政策評価制度に関する見直しを着実に進めるべく、「政策評価に関する基本方針」の改定等を平成17年内に行う、などと盛り込まれた。そして、総務省における制度見直しの作業に、ドッキングする形で反映された。こうした議論は、いわば予算との関係を特に深めて議論する経済財政諮問会議が、見直しの関係部分を深掘りする形となり、いわば「見直しのベストミックス」といってもよい形で委員会との役割分担がなされた。¹⁰ なお、この点は、本稿「はじめに」で掲げた論点③との関連で押さえておく必要がある。

③ 国会、総理官邸の議論

一方、国会においても、評価法施行から1年後に行われた「政策評価に関する決議」に続き、参議院行政監視委員会、さらには参議院本会議において、「政策評価制度に関する見直しの決議」が全会一致でなされるなど、政府の見直し作業に対するチェックが行われてきた。総務省においては、これらの動向を見直し作業に反映させた。

さらに、総務省における見直しの方向性の発表の後、総理官邸から、内閣総理大臣の通常国会の施政方針演説などで表明する重要方針について、各府省の評価の体系の中で明確に位置付けられていないとの問題意識が示され、平成17年8月の閣議において、小泉内閣総理大臣からこの点についての取組を徹底するよう指示がなされた。併せて麻生総務大臣も評価結果の予算要求等への反映の徹底を求める発言を行った。

④ 基本方針改定、新ガイドライン策定

総務省では、見直しの方向性を始めとするこ

れまでの政府の検討や、各方面の意見を総合的に検討し、今般の政策評価制度見直しでは、法改正までは必要ないものの、その運用面で改善を要する事項があるとの最終意思決定を行い、平成17年12月に、「政策評価に関する基本方針」（閣議決定）の改定、「政策評価の実施に関するガイドライン」（各府省政策評価連絡会議了承）の策定を行った。

(2) 制度見直しのポイント

① 重要政策に関する評価の徹底等

政策評価制度の導入当初は、行政改革の柱との位置付けもあり、政策全体を幅広く評価することに力点が置かれてきたが、評価に用いられる各種資源等には限りがあり、評価作業への負担感が相当大きい中で、力点を置いて取り組むべき評価活動に注力するためには、政策評価の重点化・効率化を図ることが求められるため、以下のとおり取り組むこととされた。なお、この点は、本稿「はじめに」で掲げた論点④との関連で押さえておく必要がある。

ア 内閣の重要政策に関する政策の評価

内閣総理大臣が毎年1月に召集される通常国会で行う施政方針演説等で示された内閣の重要政策について、その目標や方針を明確にした政策体系を各府省において整備し、適時的確に評価を行うようにする。

イ 政策評価の重点化

各府省において、各年度の重要政策への評価に重点を置くなど、政策評価の重点化を図ることとする。また、実績評価方式を用いて評価を実施する際には、例えば、あらかじめ設定した目標の達成度に関して毎年度実績の測定を行うこととし、その上で施策を構成する事務事業レベルに問題があるものなどについては必要に応じて掘り下げた分析を行ったり、一定期間経過後に総括的な評価を行うなど、業務量・緊急性等を勘案しつつ効率的かつ柔軟に対応することとする。

ウ 政府全体の政策の統一性・総合性を確保するための評価

関係府省は、複数府省が関係する政策の企画立案の際、政策の達成目標、目標期間、達

成目標と政策手段との関係などの明確化を図ることとする。総務省は、評価のテーマ選定に当たっては、引き続き、国民の視点に立って評価の必要性が高い分野を適切に取り上げるとともに、関係施策が極めて多岐にわたっている政策については、評価の結果を政策に適切に反映するために合理的と認められる単位により評価していくこととする。

エ 規制の事前評価の義務付けに向けた取組

規制影響分析は、「規制改革・民間開放推進3か年計画」（平成16年3月19日閣議決定）に基づき、各府省において平成16年度から試行的に実施されたが、評価法の枠組みの下で事前評価を義務付ける対象の合理的な範囲や評価を実施する時点など早期義務付けに向けた具体的な枠組みの検討を行うこととする。（その後、平成19年10月から、法律又は政令によって規制の新設又は改廃を行おうとする際、事前評価の実施が義務付けられた。）

② 評価結果の予算要求等政策への反映

政策評価は、評価のための評価にとどまってはならず、評価結果が政策に反映されて初めて意味を持つものである。このため、政策評価の結果の政策への反映を徹底するため、次のとおり取り組む。なお、この点は、本稿「はじめに」で掲げた論点①との関連で押さえておく必要がある。

ア 政策評価と予算・決算との連携の強化

「政策評価－予算－決算の連携強化を図るために、今後、施策等を対象とした評価の重要性が増すと考えられ、このことを踏まえて評価に取り組み、評価結果の予算要求等への反映の徹底を図ることとする。

イ 政策の体系化

各府省において政策体系（「政策－施策－事務事業」などの体系）をあらかじめ明らかにし、政策全体における施策等の位置付けや、施策を構成する主要な事務事業など政策の構成要素を明示した上で、政策の評価を行うこととする。

ウ 政策評価に基づいた政策の企画立案の徹底 省議などにおいて重要な政策決定が行われ

る際は、できる限り評価結果に基づいた議論が行われるようにし、予算要求などの際には政策評価担当組織と予算等取りまとめ部局が合同ヒアリングを行うようにするなど一層連携を強化することとする。

エ 事前評価の的確な実施と政策の改善・見直しへの反映

事前評価については、新規に開始しようとする事業等について、施策等との関係を明らかにした上で、必要に応じて関連する施策の事後評価の結果を踏まえ、積極的に実施することとする。その際、得ようとする効果や事後的な評価方法等を明らかにして必要に応じて事後評価など事後の検証を行うこととする。

オ 予算要求等政策への反映の説明責任の徹底 各府省において、評価結果に基づいた政策への反映の方向性を評価書に記載したり、評価結果を予算要求等にどのように反映したかのつながりを適期に明らかにするなどにより説明責任を徹底することとする。

カ 制度見直しの際などの総合的な評価

制度の改正など政策の大幅な見直しの際には、中長期的な観点から検討する必要性も高く、今後、政策を様々な角度から分析して問題点を把握するとともにその原因を分析することに適した総合評価方式などを用いて総合的に評価することとする。その際、審議会等の答申や白書等による分析結果を評価に積極的に活用することも検討することとする。なお、この点は、本稿「はじめに」で掲げた論点⑤との関連で押さえておく必要がある。

③ 評価の客観性の確保

政策評価制度は、自己評価原則を採っているが、「お手盛り評価」との批判を招かないためにも、評価の客観性を確保するシステムが不可欠である。この点は、制度見直しでも大きな論点となった。したがって、その一層の向上を図るため、以下のとおり取り組むこととする。

ア 達成目標の明示への取組

実績評価方式を用いた評価においては、まず、適切な目標や指標を設定することが肝要であり、目標設定に当たって、諸外国の事例

の調査研究の成果や各府省の取組を参考にしつつ、できる限り達成目標の定量化に取り組むこととする。

イ 政策のコスト・効果の把握

評価の客観性を確保するため、政策のコスト・効果をできる限り定量的に把握できるよう、総務省及び各府省において諸外国の事例の調査研究などの実施に努め、その結果を政府全体として共有することとする。

ウ 学識経験者の知見の活用

計画の策定など評価活動全般にわたるものについては学識経験者から成る会議を活用したり、政策の分野に応じてその専門的知見を有する学識経験者の意見を個別、具体的に聴くなど、学識経験者の知見を適切な形で一層有効に活用することとする。

エ 総務省による客観性担保評価活動

総務省による客観性担保評価活動については、各府省が行った政策評価全般を対象としており、評価の実施手続における客観性・厳格性の達成水準等に関する審査が中心となっていたが、審査対象を重点化するとともに、評価の内容面も視野に入れて活動を行うこととする。

オ 外部からの検証可能性の確保

評価結果の基礎となるデータなど評価に関する情報について、入手しやすさに配慮した形で公表することを徹底することにより、評価結果に対する外部からの検証可能性を確保し、必要に応じ、外部の意見を吸収することに努めることとする。

④ 国民への説明責任の徹底

政策評価制度の目的の一つは、行政の国民に対する説明責任の徹底であり、このためには、各府省における評価の個々の取組と、制度官庁である総務省における全体的取組の双方が相まって効果を上げることが必要であり、次のとおり取り組むこととする。

ア 国民への分かりやすさの確保、評価書・要旨の標準化

評価方式の特性を勘案して、法定記載事項が明確に記載されたものとなるよう標準的な様

式を提示し、その際、法定記載事項ごとに具体的にどのような内容を記載すべきかも明示することとする。

また、評価書の要旨は基本的に評価書ごとの概要が国民に分かりやすいものとなるように作成することを徹底し、要旨への必要的記載事項については、評価書の記載事項に沿って評価として備えるべき要素を横断的に検証できるよう、できる限り標準化する。

イ 国民的議論の活性化

国民による様々な評価や判断に資するよう政策に関する情報を提供することが政策評価の目的の一つであり、政府広報など各種メディアを通じて政策評価制度に関する情報発信を展開することとする。

⑤ 政策評価の基盤強化

最後に、政策評価制度全般を下支えし、見直しを実効あるものとするための基盤となる取組として、次のとおり取り組むこととする。

ア 職員の意識改革

政策評価制度についての問題点を適期に把握し、改善・充実に図り、政策評価制度をより一層各府省の職員に浸透させていくこととする。

イ 総務省による各府省の取組の支援

総務省において、各府省の先進的な取組を紹介したり、各府省の要望も踏まえつつ、場合によっては共同で調査研究等を実施してその結果を各府省に提供するほか、実践的な研修を実施するなど、各府省の政策評価作業に資するための活動を一層積極的に行うこととする。

ウ 地方公共団体との連携

国の政策評価を行うに当たって、必要に応じて地方公共団体の協力を得たり、地方公共団体における評価の取組を参考にするために、地方公共団体との連携に一層積極的に取り組むこととする。

(3) 制度見直しの具体的成果

「政策評価に関する基本方針の改定」、「政策評価の実施に関するガイドラインの策定を受け、各府

省は平成17年度内にそれぞれの基本計画、実施計画に反映した上で、平成18年度から、新たな枠組みの下で評価を実施することになった。筆者は、かつて見直し後の新たな枠組みによる取組を「評価新時代」と位置づけたが、今日に至るまで制度の基本的な方向は見直しの際の考え方を踏まえており、この認識は維持されていると考える。¹¹

制度導入から8年経った当時は、毎年度、各府省トータルで通常約4,000件の政策評価の実施を確保するなど制度の定着が図られた。また、評価結果は、予算編成や政策の企画立案に反映されており、公共事業については、政策評価法施行後7年間の累計で、227事業、総事業費約3.9兆円の事業の休廃止につながった。さらに、実績評価における目標の達成水準が数値化等により具体的に特定されている政策の割合も、平成14年度の34%から、21年度は82.5%と増加を見た。

また、内閣の重要政策についての評価の推進については、「経済財政改革の基本方針 2007」により、経済財政諮問会議と総務省・各府省が連携して、内閣の重要政策について評価に取り組むこととされた。

4. 事業仕分けを受けた抜本的機能強化

(1) 事業仕分け

平成21年9月、鳩山内閣の発足に伴い、内閣府に行政刷新会議が設置され、同年11月、平成22年度予算編成において、透明性を確保しながら、予算を見直すことができる有効な手法として、「事業仕分け」を実施した。

その対象事業の一つとして、「政策評価、行政評価・監視」が取り上げられることが決まり、11月13日の同会議の第1ワーキンググループにおいて、事業仕分けが行われた。

その際、様々な意見が出されたが、「とりまとめコメント」としては、抜本的な機能強化を行うこととされた。

一方で、「前向きな期待の声が出ていることを重く受け止め、これまで行ってきた評価等の業務を自己評価し、こうした期待に応えるための具体策を打ち出していきたい。」とも付言された。

事業仕分けで強化の方向が示されることは、異例の結果とすることができ、それだけ政策評価制度に対する期待が大きいことの表れであったと考えられる。

(2) 行政評価機能の抜本的強化ビジョン

行政刷新会議の事業仕分けにおける「抜本的な機能強化」との評価結果を受けて、直ちに行政評価局として具体策の検討に入り、平成22年1月12日、「行政評価機能の抜本的強化ビジョン」を取りまとめ、同日の閣僚懇談会及び第5回行政刷新会議において原口総務大臣が説明し、公表した。

同ビジョンにおいては、事業仕分けで対象とされた政策評価推進機能及び行政評価局調査機能以外に、行政相談機能及び独立行政法人評価機能も関連するものとして一体的に取り上げることとし、行政評価機能のそれぞれについて、見直しの方向性とそのための主な方策が示された。

政策評価推進機能については、「各省の情報公開徹底促進」という見直しの方向性の下、「情報公開に関するガイドライン新設」を行うこと、その他、成果志向の目標設定の推進、租税特別措置等に係る事前評価の拡充、予算編成に資する政策評価の推進などが盛り込まれた。

行政評価局調査機能については、「調査の拡充」という見直しの方向性の下、オープンな議論を経て行政評価等プログラムを決定することや、閣議等の場を活用した調査の実効性確保、「機動調査チーム」を設置することなどが盛り込まれた。その後、報道機関に対する完全公開の下、総務大臣主催の検討会を2回にわたり開催し、その議論内容は、22年4月、「行政評価機能の抜本的強化方策」としてとりまとめ、行政評価局の業務運営方針たる「行政評価等プログラム」（平成22年4月13日総務大臣決定）に盛り込まれた。

(3) 行政事業レビューとの連携と目標管理型の政策評価の導入

事業仕分けを各府省で行うこととされた行政事業レビューと連携しつつ、目標管理型の政策評価について、メリハリのある分かりやすい評価を実現するため、平成24年度からの取組についての標準的な指針を策定し、具体的には以下の取組を踏

まえた政策評価を実施することとされた。

① 事前分析表の導入

事前に施策目標を公表するとともにその達成手段（事務事業）との関係（政策体系）を整理する。

② 評価書の標準様式の導入等

各行政機関共通の標準的な様式により、統一性・一覧性を確保する。また、評価に当たり、行政事業レビューの情報を活用するなど、行政事業レビューとの連携を確保する。

なお、こうした取組の下で、政策評価ポータルサイト¹²が開設され、各府省の「政策体系」、「評価情報」及び「概算要求反映状況」を関連付けて一か所で把握できるようになったことは、評価の外部検証性の確保にとって重要な意義を有すると考えられる。

(4) 経済財政諮問会議における実効性の高いPDCAサイクルの確立

平成24年12月に発足した自公政権において復活した経済財政諮問会議において、政府内の実効性あるPDCAサイクルの確立の観点から、政策評価が再びテーマとして浮上した。

昨年12月の政権交代後、行政事業レビューの取扱いの検討や、総務大臣の制度に対する考え方を踏まえ、より実効性を高める観点からの政策評価制度の見直しが、現在まさに進行中である。なお、この点は、本稿「はじめに」で掲げた論点⑤との関連で押さえておく必要がある。

5. 政策評価制度の今後の展望

最後に、冒頭で紹介した政策評価制度をめぐる論点について、これまで振り返ってきた制度導入以来の経緯の中に見出された原因を踏まえ、その改善方策など今後の展望を示すこととしたい。

(1) 評価方式をめぐる論点

評価方式について特定の方式を前提として全体的な批判を行う場合があり、混乱・誤解を招きやすいとの論点については、我が国の政策評価制度は、制度導入時点の検討を踏まえ、多元的な評価

方式を許容しており、また、それに適した評価手法としてパフォーマンスメジャメント、費用効果分析、プロセスアナリシスなど様々なものが用いられており、議論の前提として、どのような評価方式、評価手法を取り上げているか整理することが必要である。

なお、評価基準については、目下、共通的な評価結果の表示方法を全政府的に導入することが検討されている。

(2) 評価結果の活用をめぐる論点

評価結果の活用について、評価結果をもって自動的に政策判断ができるとの過大な期待を持って議論される場合があるとの点については、政策評価は、あくまで政策の改善という意味決定のための判断材料を提供するのが本来の役割であり、政策の意思決定は、各種の意見要望や政治的判断など様々な要素を考慮してなされるものであることを押さえた上で評価結果の活用を論ずることが必要である。

(3) 評価目的をめぐる論点

評価目的について、予算の削減につながらなければ意味がないとの議論については、政策評価は政策の改善に資する情報を提供することを趣旨としており、その結果は、政策の廃止・縮小に結び付く場合もあれば、充実につながる場合もあるのであって、制度目的の全体を視野に入れて議論する必要がある。

(4) 評価対象をめぐる論点

評価対象について制度導入当初はできる限りカバーすることを重視し、かえって総花的な評価になったことについては、目下、政府内で毎年の評価の対象を重点化し、評価内容を深掘りすることが検討されている。

(5) 評価時期をめぐる論点

評価時期について、政策の実施が終わってから評価が始まるとの考え方が一般的である点については、政策の企画立案から実施、評価に至る過程は、PDCA（plan-do-check-action）やPDS（plan-do-see）などのサイクルで説明されるものである

が、政策の企画立案に際しては事前評価が必要であること、政策の実施の最中にもプロセスチェックがあることなど、これらの過程は実際には渾然一体となって進んでおり、日常的な業務遂行の中に評価作業が内在することを視野に入れる必要がある。

むすびにかえて

政策評価制度は、スタートから10年以上が経過した。この10年間で、政策評価制度をめぐる環境は、各府省や総務省による着実な取組により大いに進展し、国の行政運営の中に、定着するに至ったことは大きな成果であったと言える。

この間、総務省の諮問機関である政策評価・独立行政法人評価委員会、さらには、国会や経済財政諮問会議、官邸など各方面からの問題提起を受け、一つ一つ課題をクリアし、発展過程を迎えた。こうした流れは、2度にわたる政権交代においても何ら変わることはなく、むしろ強化されていると言える。

制度見直しを経た「評価新時代」においては、従来とは一線を画し、評価対象の選定、内容の質の向上、結果の政策への反映といった評価の各ステージでこれまでよりバージョンアップした取組が求められたが、政府の取組に対する各方面からのチェックもより厳しくなることが想定され、説明責任を果たす必要性も一層増大すると考えられる。

しかし、制度導入に向けてはゼロベースからの検討を余儀なくされたことを考えれば、中長期的にみて政策評価制度は着実に進化・発展を遂げてきており、相当なノウハウが蓄積されてきていると考える。今後も、政府一丸となって、「評価新時代」にふさわしい取組が実践されるよう願ってや

まない。

注記

- 1 行政管理研究センター（2007）p5
- 2 行政管理研究センター（2007）p32
- 3 行政管理研究センター（2001）p52
- 4 行政管理研究センター（2007）p99
- 5 行政管理研究センター（2007）p50
- 6 政策評価の手法等に関する研究会（2000年）p26
- 7 行政管理研究センター（2008）p9
- 8 政策評価制度の法制化に関する研究会（2000年）p7
- 9 行政管理研究センター（2007）p102
- 10 行政管理研究センター（2007）p103
- 11 行政管理研究センター（2007）p123
- 12 http://www.soumu.go.jp/main_sosiki/hyouka/seisaku_n/portal/index.html を参照

参考文献

- 行政改革会議（1996）『最終報告』
 政策評価の手法等に関する研究会
 （2000）『政策評価制度の在り方に関する最終報告』
 政策評価の法制化に関する研究会
 （2000）『政策評価制度の法制化に関する研究会報告』
 行政管理研究センター（2001）『政策評価ガイドブック—政策評価制度の導入と政策評価手法等研究会』
 ぎょうせい
 行政管理研究センター（2007）『政策評価ハンドブック—評価新時代の到来』 ぎょうせい
 行政管理研究センター（2008）『詳解政策評価ガイドブック - 法律、基本方針、ガイドラインの総合解説』
 ぎょうせい

（2013.9.19受理）

図1 政策評価制度に関する経緯

平成9年12月	行政改革会議最終報告－政策評価の導入を提言－
平成10年6月	中央省庁等改革基本法成立－中央省庁等改革の基本方針として政策評価機能の強化が盛り込まれる－
9月	中央省庁等改革に係る立案方針（中央省庁等改革推進本部決定）－政策評価機能の充実強化を図るための立案方針を提示－
平成11年1月	中央省庁等改革に係る大綱（中央省庁等改革推進本部決定）－政策評価に関する大綱を提示－
4月	中央省庁等改革の推進に関する方針（中央省庁等改革推進本部決定）－政策評価の枠組みを決定－
5月	総務庁行政監察局に政策評価等推進準備室を設置 各省庁政策評価準備連絡会議発足－政策評価の実施方法等に関する標準的ガイドライン案の策定作業に着手－
6月	中央省庁等改革関連十七法律案に対する衆議院・行政改革に関する特別委員会における附帯決議－政策評価制度の法制化につき早急の検討を指示－
7月	中央省庁等改革関連十七法律案に対する参議院・行財政改革・税制等に関する特別委員会における附帯決議－政策評価制度の法制化について早急の検討を指示－ 国家行政組織法一部改正法、内閣府設置法、総務省設置法成立－政策評価の根拠規定が盛り込まれる－
8月	各省庁政策評価準備連絡会議において、「標準的ガイドライン案に盛り込むべき事項（検討案）」了承 政策評価の手法等に関する研究会・初会合
9月	「標準的ガイドライン案の検討方向案」を取りまとめ、中央省庁等改革推進本部第15回顧問会議に報告
平成12年2月	政策評価の手法等に関する研究会が「政策評価の導入に向けた意見・論点の中間整理」を公表
4月	「政策評価に関する標準的ガイドラインの試案（概要・全文）」を取りまとめ、中央省庁等改革推進本部第17回顧問会議に報告
5月	新たに総務省に設置される政策評価・独立行政法人評価委員会の組織、所掌事務等を規定した政策評価・独立行政法人評価委員会令が閣議決定
6月	政策評価の手法等に関する研究会が「政策評価の導入に向けた中間まとめ」を公表
7月	各省庁政策評価準備連絡会議において「政策評価に関する標準的ガイドラインの案（概要・全文）」を了承し、公表
8月	総務庁行政監察局に政策評価制度法制化担当室を設置
9月	「政策評価に関する標準的ガイドラインの案」に関する意見募集の実施 政策評価制度の法制化に関する研究会・初会合
11月	「政策評価に関する標準的ガイドラインの案」に関する意見募集結果公表

12月	行政改革大綱－政策評価制度の法制化及び第151回国会への法案の提出を明記－ 各省庁政策評価準備連絡会議において「政策評価に関する標準的ガイドライン（案）（概要・全文）」を了承し、公表 政策評価の手法等に関する研究会が「政策評価制度の在り方に関する最終報告（概要・全文）」を公表 政策評価制度の法制化に関する研究会報告（概要・全文）」を公表
平成13年1月	政策評価各府省連絡会議発足 同会議において「政策評価に関する標準的ガイドライン（概要・全文）」を了承し、公表 政策評価・独立行政法人評価委員会が発足
3月	「行政機関が行う政策の評価に関する法律案」の閣議決定、国会提出
6月	「行政機関が行う政策の評価に関する法律案」、衆議院で一部修正、可決 参議院で可決・成立 「行政機関が行う政策の評価に関する法律（概要・全文）」の公布
9月	行政機関が行う政策の評価に関する法律第5条第4項の審議会等で政令で定めるものを定める政令の閣議決定、施行－政府全体の政策評価を計画的かつ着実に推進していくための指針となる政策評価に関する基本方針の案の作成に際し意見を聴く審議会等として、政策評価・独立行政法人評価委員会を指定－ 総務省行政評価局に行政機関が行う政策の評価に関する法律施行準備室を設置－政策評価に関する基本方針、事前評価の義務付けの対象とする政策等を定める法施行令の策定作業に着手－
12月	政策評価・独立行政法人評価委員会において「政策評価に関する基本方針案」の諮問・答申 「政策評価に関する基本方針」閣議決定
平成14年3月	「行政機関が行う政策の評価に関する法律第5条第4項の審議会等で政令で定めるものを定める政令の一部を改正する政令」（法施行令）が閣議決定－事後評価が義務付けられる政策の要件となる政策決定からの未着手又は未了の期間を決定するとともに、事前評価が義務付けられる政策を決定－
4月	「行政機関が行う政策の評価に関する法律」及び「行政機関が行う政策の評価に関する法律施行令」が施行－各行政機関は、法律、政令及び基本方針に基づき、基本計画及び実施計画を策定して政策評価を実施－
平成15年6月	政策評価の円滑かつ効果的な実施について（局長通知）
9月	規制に関する政策評価の手法に関する研究会を開催（16年7月まで6回）
平成16年4月	政策評価分科会の当面の活動の重点～政策評価制度の発展に向けて～－政策評価制度の見直しに向けた取組を政策評価分科会が表明－
7月	規制に関する政策評価の手法に関する調査研究報告書を公表
10月	「規制影響分析（RIA）の試行的実施に関する実施要領」に基づき規制影響分析の試行的実施を開始
12月	政策評価法附則第2条に基づき、法施行後3年が経過したことを受けて、政策評価制度に関する見直しの論点整理を公表

政策評価制度10年の軌跡
～制度導入以降の省察と今後の展望～

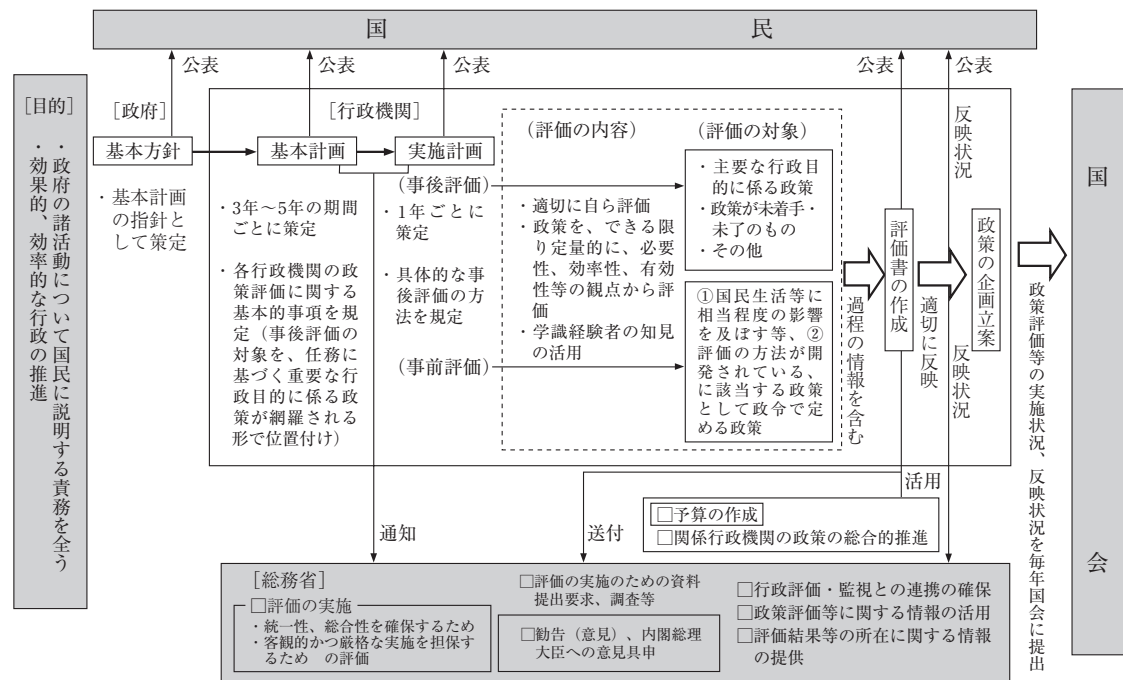
17

平成17年6月	政策評価制度に関する見直しの方向性を公表 「政策評価制度に関する見直しの方向性」の発表に当たって（政策評価・独立行政評価委員会委員長、政策評価分科会長談話）
9月	規制の政策評価に関する研究会を開催（平成19年10月まで計11回開催）
11月	規制の政策評価に関する研究会中間報告を公表
12月	政策評価に関する基本方針の改定について（答申） 「政策評価に関する基本方針の改定案」の答申に当たって（談話） 政策評価に関する基本方針の改定について 政策評価の実施に関するガイドライン策定について 行政改革の重要方針を閣議決定－規制の事前評価の義務付けや、政策評価の重点化・効率化を推進することを閣議決定－
平成18年3月	規制改革・民間開放推進3か年計画再改定－規制影響分析の義務付けについて閣議決定－
6月	簡素で効率的な政府を実現するための行政改革の推進に関する法律制定－内閣の重要政策に係る政策評価の重点的かつ効率的な実施を推進することを規定－
平成19年3月	「行政機関が行う政策の評価に関する法律施行令」の一部改正、「政策評価に関する基本方針」の一部変更－事前評価の義務付けに規制を追加－
6月	経済財政改革の基本方針2007において、経済財政諮問会議と総務省・各府省の政策評価に関する連携を強化することを明記 政策評価の重要対象分野に関する取組を開始
8月	「規制の事前評価の実施に関するガイドライン」の策定、「行政機関が行う政策の評価に関する法律施行規則」の制定
9月	規制の政策評価に関する研究会最終報告を公表
10月	規制の事前評価の義務付けを開始
平成22年4月	規制の事前評価における競争状況への影響の把握・分析の試行的実施を開始行政機関が行う政策の評価に関する法律施行令の一部を改正する政令案等に係る意見募集を実施
5月	政策評価に関する基本方針の一部変更ほか政策評価の機能強化方策について、政策評価・独立行政評価委員会への諮問・答申 行政機関が行う政策の評価に関する法律施行令の一部を改正する政令案等に係る意見募集の結果を公表 「行政機関が行う政策の評価に関する法律施行令」の一部改正 「政策評価に関する基本方針」の一部変更 「政策評価に関する情報の公表に関するガイドライン」及び「租税特別措置等に係る政策評価の実施に関するガイドライン」を策定
平成23年4月	目標管理型の政策評価の改善方策に係る試行的取組を開始
平成24年3月	「政策評価の実施に関するガイドライン」の一部改正 「目標管理型の政策評価の改善方策に係る取組について」を策定

（出所）総務省ウェブサイト

図2 行政機関が行う政策の評価に関する法律（骨格）

行政機関が行う政策の評価に関する法律の概要



(出所) 総務省ウェブサイト

Ten-year trajectory of the Policy Evaluation System ～Retrospective and prospective analysis of the System～

Seiichi Arai

Administrative Evaluation Bureau, Ministry of Internal Affairs and Communications
sarai@soumu.go.jp

Abstract

In Japan, the Policy Evaluation System was introduced to the administration in 2001. This system was launched to improve the administration that traditionally focused on establishing laws or getting resources, when a standard guideline on policy evaluation was formulated by the Ministry of Internal Affairs and Communications.

When the system was finally legislated, it became firmly established in the policy management cycle (the PDCA cycle) of each ministry. In 2005, three years after the enforcement of the Government Policy Evaluation Act, the system was totally reviewed. It was later on decided that the functions of the Policy Evaluation System should be drastically enhanced as a result of the budget screening process in 2009.

However, there are several arguments concerning the Policy Evaluation System.

In my paper, I would like to find the cause of these arguments by looking back on the progress that the system has made and unveil its visions of the future.

Keywords

Policy Evaluation System, Policy Management Cycle, Total review of the Policy Evaluation System,
Enhancement of functions of the Policy Evaluation System

【研究論文】

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～

小野 達也

鳥取大学地域学部

onotatsu@rstu.jp

要 約

2001年に導入された政府の政策評価制度では、一般の政策に係る事後評価の大半は「実績評価」として行われ、2011年度からは「目標管理型の政策評価」として、この実績評価が行政事業レビューと連携して実施されている。実績評価方式は欧米諸国と同様の実績測定（パフォーマンス・メジャーメント）を想定したものであり、行政事業レビューも実績測定の性格をあわせ持つ。しかし、実績測定の中心となる機能、すなわち実績を測定するという機能がこの両者にどこまで備わっているのか、制度と運用の両面について考察すれば、多くの課題が浮き彫りになる。今後、実績測定をより本格的に実行しようとするのであれば、これらへの対策が急務である。

キーワード

実績測定（パフォーマンス・メジャーメント）、実績評価、目標管理型政策評価、行政事業レビュー

はじめに

2001年1月6日に導入され、2002年4月1日には「行政機関が行う政策の評価に関する法律」（以下「評価法」と呼ぶ）が施行された政府の政策評価制度では、欧米諸国の実績測定（パフォーマンス・メジャーメント）¹を想定した「実績評価」が標準的な評価方式の1つとして示され、一般の政策（評価法によって評価が義務付けられた政策—当初は公共事業・研究開発・ODAの3分野—以外の政策）に係る事後評価の大半がこの方式によって行われることとなった。

また、民主党政権下の2011年度からこの実績評価は「目標管理型の政策評価」として、政策評価

制度とは別の枠組み²で導入された行政事業レビュー（実績測定の性格を明確にもつ）と連携する形で実施されている。この両者の連携は2012年12月からの自民党政権下においても継続しており、また今後は政策評価と行政事業レビューの連携強化や目標管理型評価の重点化・質向上のための検討がなされることとなっており、実績測定型の評価（実績評価・目標管理型政策評価、行政事業レビューなど実績測定の機能を有する評価制度・システムを以下では「実績測定型評価」と総称する）が今後も重要な役割を担うことが確実である。

実績測定・実績測定型評価には長い歴史があり、様々な経緯を経て、現在も世界の多くの国々において、様々な形で行われているが、その中心

は実績の測定である。測定とは量を測ることであるから、実績測定とは、評価対象のプログラム（狭義政策・施策・事業を本稿ではプログラムと総称する）の実績に関する量に、ある手順に従って数値を割り当てることである。測定結果の数字、すなわち実績指標³（の値）をどのように活用するかは、当然評価の目的や制度・体制によって異なるだろうが、測定自体の品質が一定程度確保されていることが、何をやるにしても大前提となることはいうまでもない。

本稿はこのような観点から、府省の実績測定型評価における計量・計数の課題を論ずるものである。本稿の構成は次の通りである。第1節では、後段の議論のために、実績測定とはそもそも何か、評価手法・制度としての特徴やその役割と課題を巡る議論を概観し、第2節で日本の府省が行う実績測定型の評価に係る制度を確認する。それらの準備を踏まえ、第3節において計量・計数に関わる一連の局面について、求められる条件と現状を対比させながら課題を明らかにする。最後の第4節では本稿の観察と考察を総括し、得られる示唆について述べる。

政策評価制度の実績評価方式については、実績測定（型評価）の本来の役割・機能という観点（山谷 2006, 2012）、また府省における現実の政策過程との関係という観点（田辺 2001, 2005, 2006, 2007）などから分析・批判・提言がなされているが、本稿の目的は実績測定の基本作業・基本ツールである測定・指標の現状について点検・考察することにある。

府省の実績測定型評価の測定部分の一斉点検は、管見では総務省行政評価局の客観性担保評価活動（後述）のみである⁴。また、指標による測定に関する理論的考察は、指標の妥当性・信頼性に関しては後述するように多くの文献が論じている一方、指標設定後の比較について評価研究の分野で体系的に論じたものは、管見によれば筆者（小野2003, 2008, 2011a, 2011b）を除いて見当たらない。ただし、個々の事項自体は何れも社会統計学や社会指標研究の分野で考察されている。

1. 実績測定とは

実績測定とは何か。米国GAOは「一般に（commonly）」次のように定義されるとしている（U.S.GAO1992）－実績測定とは、インプット・作業（活動）・アウトプット・アウトカム・効率（生産性）などの一連のデータを定期的に収集し報告すること。これが実績測定の核心であるが、公共部門の制度・システムとしての特徴を敷衍して述べれば、実績測定とは、継続中（又は終了後）の行政活動の集合を対象として、その実績を測定・解釈（場合により分析・評価）するもので、しばしば事前の段階での将来予測や目標設定を含む過程であり、特徴として、広範な評価対象について恒常的に実施すること、評価対象に対して統一的な測定・評価基準を用いること、特定の目標に対する達成度に重点を置く場合も多いこと、組織全体のパフォーマンスに強い関心があること、評価というよりも測定・報告が中心である場合も多く、広く浅いモニタリングによる早期の警告とアカウンタビリティ遂行が重要な役割であること、通常は行政内部で行われることなどの諸点を挙げるができる。

実績測定の起源を遡ればその歴史は長いが、世界中の公共部門に急速に普及したのはNPMの時代になってからである（今日に至る歴史的経緯については古川 2002、山谷 2012、Van Dooren et al. 2010、Hatry 2013等を参照）。プログラムを正面から掘り下げて厳密に評価しようというプログラム評価と比べれば、実績測定は表面的で記述的な説明に過ぎないことは明らかであるが、プログラムや機関のパフォーマンスの現状を包括的に明らかにできることから、公共部門改革の潮流において熱狂とともに受け入れられることとなった（Greiner 1996、Poister 2004）。

実績測定とプログラム評価はそもそも目的が異なるが、プログラムに関する様々なパフォーマンスをモニタリングするという手法自体はプログラム評価の1つの過程に相当することもあり、プログラム評価を重んじる立場からは、実績測定についてかつて多くの批判的言及・考察がなされ様々な議論が行われたが、今日では両者は互いに補完するものであるという認識が概ね共有されるに至

っている (Nielsen&Hunter 2013、Hatry 2013)。

ただし、実績測定を実践しようとすれば、実績指標という簡素なツールでどこまで厳密さを追求すべきか、その評価結果を意思決定にどこまで直結すべきか、様々な分野のプログラムにどこまで統一性を求めるか、効率化のための改革をどこまで内部評価に託せるのか、といったいわば単純な解のない難題を常に抱えることが避けられない。したがって、実績測定の実践においては試行錯誤が宿命づけられているともいえ、多くの文献が実績測定のパラドクスやジレンマを論ずるとともに、様々な対策の提案が続いている (Nielsen& Hunter 2013、Van Dooren et al.2010、De Bruijn 2007、Poister 2004、Newcomer 1997、Osborn& Gaebler 1992等)。

2. 府省における実績測定の制度

(1) 政策評価制度と実績評価方式

2001年12月28日、翌年4月の評価法施行を前に閣議決定された「政策評価に関する基本方針」は、政策評価の方式について「政策の特性等に応じて合目的に、「事業評価方式」、「実績評価方式」及び「総合評価方式」(別紙)やこれらの主要な要素を組み合わせた一貫した仕組みなど、適切な方式を用いるものとする。」とし、別紙では実績評価方式を「政策を決定した後に、政策の不断の見直しや改善に資する見地から、政策の目的と手段の対応関係を明示しつつ、あらかじめ政策効果に着目した達成すべき目標を設定し、これに対する実績を定期的・継続的に測定するとともに、目標期間が終了した時点で目標期間全体における取組や最終的な実績等を総括し、目標の達成度合いについて評価する方式」と定義している。これは実績測定、それもアウトカムを重視するその定義そのものである。

政策評価制度の導入後間もない2001年1月15日に決定された「政策評価に関する標準的ガイドライン」では、「標準的な三つの評価の方式」の1つとしての実績評価が詳しく説明されている。ここでは原則として施策レベルでアウトカムの基本目標を具体的な達成水準として示すこと、それが難

しい基本目標は関連した測定可能なアウトカム・アウトプット指標(可能な限り客観的に達成度を測定できるような定量的又は定性的な指標)を用いて達成目標を具体的に設定すること、目標や指標の設定に当たっては設定の考え方等を明らかにすること、アウトカム指標については外部要因を可能な限りあらかじめ明らかにすること、などが述べられている。

また、ガイドライン策定作業と並行して設置された研究会の報告書(政策評価の手法等に関する研究会 2000)には、米国や英国などの実績測定が、実績評価方式のモデルであることが述べられている。

評価法によって評価が義務付けられていない一般の政策の事後評価は、制度導入以来その大半(年間約3百～4百本、施策レベルではほとんど)が実績評価方式により、即ち実績測定として行われている。

(2) 目標管理型政策評価と行政事業レビュー

民主党政権下の2011年度には、「目標管理型の政策評価」(実績評価方式を用いた政策評価及びあらかじめ設定された目標の達成度合いについて評価する内容を含む、いわゆる「施策」レベルの政策の事後評価)(総務省 2011)の改善方策が試行的に取り組みされた。そのねらいは、行政事業レビューと連携しつつ、各行政機関の政策のミッションの明確化・体系化及びメリハリのある分かりやすい政策評価を実現することである。

2012年度には事前分析表、評価書の基本様式(各府省においてカスタマイズ可能)が導入され、目標管理型の改善方策が本格的に実施された。事前分析表は測定指標の選定理由及び目標値(水準・目標年度)の設定の根拠を明記する欄があるほか、施策目標とその達成手段たる事務事業の関係(政策体系)を整理するものである。評価書の基本様式は、中央に測定指標欄(基準値・実績値・目標値(目標年度及び各年度)などを記載)が据えられ、目標の達成状況等を記述する欄がある。目標管理型評価はシンプルな形式で実績測定を統一的に徹底しようとするものと言える。

行政事業レビューとは、各府省がムダの削減と効率的・効果的な事業の実現を目指し、自らの事業を点検するもので、2010年度に当時の行政刷新

会議の指揮により導入された。2012年度には6千を超える全事業を対象に全府省統一のレビューシートを用いて行われている。

レビューシートでは、事業に関する基礎情報に続いて成果（アウトカム）指標の名称・実績（3年度分）・目標値・達成度（3年度分）、活動（アウトプット）指標の名称・実績（3年度分）・当初見込み（4年度分）、単位当たりコストとその算出根拠を記載する欄があり、続く「事業所管部局による点検」欄では「活動実績、成果実績」関連4項目を含む13項目について○△×の3段階評価を行う。指標の記載欄は実績測定そのものであり、その前後部分を含めて実績測定型評価の典型となっており、シートのここまでの部分は自治体で広く行われている事務事業評価型評価の定型フォーマットにもかなり近い。ただし、この後に、行政事業レビューのもう1つの重点である事業予算の使い途・流れをチェックする欄が続く。

目標管理型政策評価と行政事業レビューの連携は、2012年12月からの自民党政権下においても継続しており、本稿執筆時点では、全政府的なPDCAサイクルの重視という方針のもと、2014年度以降の政策評価と行政事業レビューの連携強化や目標管理型評価の重点化・質向上のための検討がなされることとなっている⁵。政府レベルでの政策評価重視という姿勢は明らかであり、その中で実績測定型の評価が重要な役割を担い続けることも確実である。

3. 計量・計数に関わる課題

(1) 実績測定型評価の現状—実績の測定がなされているか

評価の研究と実務に長く携わってきたホーリーは、実績測定システムが有効であるための前提条件として、①合意に達している目標や戦略がある、②測定システムの質が十分に高く①に関するパフォーマンスの評価（assess）・報告ができる、③実績情報が明確で理解しやすく意味がある、という3点を挙げている（Wholey 2007）。評価対象が定まっていなければ、そもそも評価の意味はないし、肝心の測定の品質が悪ければ、その後の議論の意義は危うい

ものとなるから、至極当然の条件である。

さて、府省の実績測定型評価の現状はどうか。本稿の執筆にあたり、筆者は2011年度に行われたプログラムを対象として2012年度に実施された目標管理型政策評価と2012年度に実施された行政事業レビューにおける指標・目標値の設定状況について調査を行った（以下「今回の調査」と呼ぶ）。

目標管理型政策評価は17機関が実施しており（16機関は実績評価方式、1機関は総合評価方式）、機関ごとに10本の評価書を系統抽出法に準じた方法で抽出した⁶。評価書の数すなわち評価対象プログラムが10本に満たない機関についてはその全数を調査対象としている。その結果17機関の合計448本の評価書のうち144本が調査対象となった。行政事業レビューは17機関が実施しており（うち16機関は目標管理型評価も実施）、目標管理型評価と同様の方法で、政府全体で6301枚のシートから163枚を抽出した。

今回の調査結果から、まずは測定の前提となる指標の設定状況をみよう。目標管理型の政策評価では、指標が全く設定されていない評価書は2機関の7本（全144本の5%）のみで、ほとんどの評価書で何らかの指標が設定されている。指標（多くは基本様式と同じ測定指標と呼ばれる）の合計は806個で、評価書1本当たり5.6個であった。機関によって政策体系のどの階層を目標管理型評価の対象とするか、1本の評価書において指標をどれだけ設定するかは区々であるため、機関ごとの評価書1本当たり平均指標数は、0.9個から17.4個までの幅がある。

ただし、評価書において指標とされているものは、すべてが数・量によって測られる定量指標ではない。実績評価方式では、定量的に把握できない場合は、数・量の測定によらない定性指標を設定する。ただし定性指標であっても進捗や目標達成状況を判断できるものであることが求められる。目標管理型評価の評価書において指標と称するもののうち、定量指標は80%、残りの20%は定性指標で、定性指標ばかりで定量指標が設定されていない評価書が15本（10%）あった。定量指標の占める割合は機関によって差があり、6機関で全指標が定量指標である一方、定量指標の割合が

3割台の機関が2ある。

行政事業レビューでは、成果指標・活動指標の何れも設定されていないシートが163枚のうち31枚（19%）あった（指標欄に何らかの記述があっても指標の体を成していないものは指標として数えず）。一方、成果指標または活動指標を複数設定しているシートは29枚（18%）で、大半は成果指標・活動指標とも1個以内を設定している状況である。なお、行政事業レビューの指標はすべて定量指標である。

評価指標が設定されていない場合は、ホーリーの掲げる前提条件以前の問題であり、実績測定の体を成していない。目標管理型評価については、指標が設定されることが大前提であり、少数ながら指標のない評価書は、評価をしていないに等しい。行政事業レビューについては、19%（15機関にまたがる31枚）という少なからぬ割合のシートが、実績測定の機能を放棄ないし拒絶していることになる。

何らかの評価指標が設定されて初めて、ホーリーの前提条件について吟味できる段階となる。ただし、定性指標については、実績がプログラムの進捗状況を明らかにし、目標が達成状況を判断できる形で設定されて初めて、本来の評価指標といえる。次項以下では、指標や目標値など実績測定型評価における測定の品質について具体的に吟味することとする。

ところで、政策評価制度の実績評価の品質については、総務省行政評価局が2002年度～2009年度の間、客観性担保評価活動として、各府省の評価の点検を行っている。そこでは次のような指摘がなされている－アウトカムに着目した目標や指標の設定がなされていない、設定された指標が政策全体をカバーしていない、設定された指標と政策目的との関係が明らかでない、設定された目標と関連する政府の計画が整合しない、指標の変化が政策効果に係る状況の変化を十分に説明していない、目標が数値化されていない、「目標年度において向上させる」という達成目標では各年の達成度合いの判断ができない、目標の達成度合いの判定の根拠や理由が明らかでない、測定指標の状況を踏まえた評価結果（分析・記述）になっていない、実績評価方式が機能しない場合は他の評価方

式や方法を検討すべき、など。

これらの指摘（毎年のように何れかの府省に対してなされている指摘事項もある）は、目標を数値化する割合が一定程度上昇するなど少なくとも一定の効果は上げている。また2011年度に導入された「目標管理型政策評価」という改善方策も、測定指標の選定理由及び目標値の設定根拠を明記するための欄や目標の達成状況等を記述する欄が基本様式として定められたことは、これまでそのような情報を示していなかった機関に一定の効果을及ぼすことも期待できる。

このように政府内でも実績測定型評価の品質向上の必要性は認識されているといえるが、以下では、指標を設定する段階、指標を使う（比べる）段階の順に、後者については、目標値を設定し達成度を明らかにする手順（目標管理型政策評価と行政事業レビューの双方において求められている）に絞って、現状を点検し、課題を明らかにする。

(2) 指標の妥当性

一般に社会科学における測定では、妥当性 validity（測定すべきものを測定していること）と信頼性 reliability（同じ事象の測定からは同じ結果が得られること）をできるだけ充たすことが求められる。社会科学においては、自然科学のように測定対象の事象とデータが密着していないため、常に測定・データの妥当性と信頼性を吟味する必要がある⁸。実績測定は、公共プログラムのアウトカム・効率などのパフォーマンスという事象を測定し、その結果を政策サイクルに反映させようというのであるから、実績指標の妥当性と信頼性は文字通りの死活問題である（Nielsen & Hunter (2013) の表現を借りれば、実績測定においてデータの品質は根源的な必要条件であり、他の課題とは別次元の課題である）。実際、実績測定の方法を解説する多くの文献も、指標・測定の妥当性・信頼性を欠くべからざる条件として述べている（Langbein 2012、Wholey 2007、Poister 2004、Rossi et al. 2004、Swiss 1995等）⁹。

とりわけ、妥当性は信頼性を問うための前提条件でもあり、それが損なわれればすべてが台無しになるゆえ、さらに別格の注意を払う必要がある。実績指標の妥当性は、2段階に分けて考えると、

その検討や設定において考慮すべき条件が明確になる。第1段階は、プログラムのロジック（投入～過程～アウトプット～直接的・中間的・最終的アウトカムの論理的関係）のどの局面を測定するかという論理に基づく指標設定に係る妥当性（本稿では「論理的妥当性」と呼ぶ）であり、第2段階は具体的に当該局面の何を測定するかという算式決定に係る妥当性（本稿では「操作的妥当性」と呼ぶ）である¹⁰。

論理的妥当性は、測定する事象とプログラムの間に因果関係があるか（評価対象プログラムのアウトカムやアウトプットなどどの局面を測定するのか）、その測定が当該プログラムの実績の測定となるのか、測定結果をもってプログラム全体の結果とみることができるか（代表性）などの観点である。つまりプログラムのロジック・モデル（ロジックを連鎖図などで表したもの）上のどのブロックを実績指標の測定対象とするかが明確であり、かつその選択が適切であることである。

操作的妥当性は、いかなるデータから指標を計算するのか、その算出方法・計算式が適切なものであるかということである。どのような業務統計や調査結果を使うのか、あるいはデータ収集のための測定やアンケートを行うのか、何れの場合にも指標設定時に設計するデータ収集対象の代表性やカバレッジ、タイミングなど多くの点が操作的妥当性に関わる。

また基礎データが決まってからは、実数をそのまま指標とするのか、比の形に加工するのか、何らかの標準化を施すのか、などの選択があり得る。要は、パフォーマンスの量（あるいは質）の大小・上昇下降などの変化が、指標の値に正しく反映されるか（もちろん単純な比例関係である必要はない）ということである。

さて、目標管理型政策評価における指標は、「政策効果に着目した達成すべき目標」という表現から明らかなように、アウトカム指標の設定を想定している。今回の調査において（定性指標を含む）全指標に占めるアウトカム指標（筆者の判断でアウトカム指標とアウトプット指標の何れかに分類した）の割合は44%であった。機関による差も大きく、アウトカム指標の割合が72%の機関がある一方、3機関では1割に満たない（ここでは

あくまでも原則通りに実施されているか否かを問うており、アウトカム指標の設定を一律に求めることの是非については後述）。

次に、アウトカム・アウトプット何れの指標であれ、その妥当性を示す情報が事前分析表の「測定指標の選定理由及び目標値（水準・目標年度）の設定の根拠」欄か、評価書・評価実施計画のどこかに示されているか否かであるが、今回の調査によれば、指標の設定理由など多少なりとも妥当性に関わる情報が何れかに記載されていると見なせる評価書は全体の53%に止まっている。

ただし、幸いなことに一部の省（農林水産省や国土交通省など）で、事前分析表の該当欄（具体的な記述がなされればこの欄は必ず一定程度大きくなる。逆にこの欄が小さいままであるのは不自然である）や参考資料、あるいは実績評価書の中でかなり詳細に評価指標と目標値の設定根拠を記述している例（論理的妥当性・操作的妥当性をどの程度満たすかも読み取ることができる）も見られる。

なお、実績測定において指標を設定する際のツールとしてHatry（2007）など多くの文献が推奨するロジック・モデルは、かつて総務省など実績評価書に記載される例が僅かにあったが、現在は皆無である。評価書に記載されていなくても、公表されていないバックヤードで指標設定のために作成されている、あるいは政策の立案段階で作成されたロジック・モデルに基づいて設定されている、という可能性についても、今回の調査で目を通した評価書等の記述からは窺えない。

行政事業レビューでは、今回の調査で調べたシートの81%で、成果指標・活動指標の少なくとも1つが設定されていたが、成果指標の設定率は44%に止まる（記載があっても明らかなアウトプット指標である場合はカウントせず）。アウトプット指標の設定率は79%であった。

シートには成果指標・活動指標ともその妥当性に係る情報を記載するための欄はなく、指標名とともに何らかの説明が書かれていたり、指標名自体にその意味がわかるような表現が含まれたりする場合が多い（成果指標では92%、活動指標では87%—筆者の判断による）ものの、指標の妥当性を示すような詳しい説明は望めない。

(3) 目標値を設定し達成度を明らかにすること

指標が設定され、その実績値が得られれば、その値が評価対象プログラムについて何を意味するのか、例えば成果をどれだけ上げたのか、順調に進んでいるか否かなどの判断をする段階となるが、この判断は専ら比較によってなされることとなる。指標の値は何かと比較することで初めて判断に資する意味をもつからだ。

この比較のなかでも、指標値の変化について掘り下げた分析を行わないことが前提の実績測定において大変有効であり、実際広く行われているのが実績値を目標値などの基準と比べる方法である。目標管理型政策評価とはまさにこの比較を中心に行われるものであり、当初からの実績評価方式も理念は同様であった。行政事業レビューも、この比較を求めている。

この比較は、シンプルな作業によってプログラムの成果や状況の良し悪しを明確に判断し、しかもその度合いを数量的に把握することができる。通常、政策・施策・事業の何れかの階層の多数のプログラムが同時にモニタリングされる実績測定型評価では、指標の目標達成度に基づいてプログラム間の比較が行われる場合も多い。

しかし、実績測定型評価では個々のプログラムの評価は比較的少数の指標によって行われる。それらの実績値や目標値という少数の数字に基づいて重要な判断をするのであれば、その比較作業には高度の慎重さが求められることとなる。妥当性を有する評価指標を設定し、また信頼性の高い測定が可能であるとして、その指標の目標値を設定し実績値と比べる作業、その目標達成度に基づいてプログラム同士を比べる作業には、この比較自体の成否を左右する重要なポイントがいくつかある¹²。これらの中には、上で言及した政策評価制度に係る総務省行政評価局の点検において指摘がなされているものもあるが、公表されている評価情報を見る限り、目標管理型政策評価・行政事業レビューともに取扱いが不十分である場合が多い。以下、6点について具体的に述べる。

①目標値の設定

指標の実績値が目標年度または各年度の目標値を達成できたか否か、どの程度達成したかを判断

するためには、当然の前提として、定量指標について明確な数値目標を設定するか、定性指標の場合は明確に判断できるような指標の定義と目標の記述がなければならない。

今回の調査では、目標管理型評価で設定された定量・定性の指標806個のうち目標達成の有無と程度が判断できるものは定量指標330個（全指標の41%、定量指標の51%）、目標達成の有無だけが判断できるもの（定量指標で「前年度以上」¹³、定性指標で「〇〇を実施」などの目標設定）は118個（全指標の15%）であった。合わせても指標の56%であり、残りの44%は「目標管理できない」指標である。ここでも機関による差は大きく、目標管理できない指標の割合は0%から100%までの幅がある。

行政事業レビューでは、成果指標について目標値の設定が求められている（活動指標は「見込み」の記載が求められる）。しかし成果指標71個のうち、目標値の設定は49個（69%）に止まる。なお、行政事業レビューでは「前年度以上」といった定性的目標は見られない。

②目標値の妥当性―目標値の性格と設定根拠

目標値の妥当性とは、目標値として掲げられた水準が、当該プログラムが本当に達成すべき水準であるか、ということである。その前提として「達成すべき」ということの意味、すなわち目標値の性格が明確でなければならない。目標値がどのような性格をもつか、どのようなスタンスで設定するかによって、目標の達成、未達の評価は当然変わってくる。

目標値の性格には、少しでも近づきたい理想的な水準、着実に達成することが期待できて一定程度満足できる水準、絶対に達成すべき義務的な水準など様々あるはずであり¹⁴、この違いにより達成度の解釈は当然別のものとなる。目標値に丁度届いた、目標値に僅かに届かなかったなど、何れの場合も目標値の性格によってその評価は全く異なるものとなろう。

また、目標値の数字自体の根拠も明らかでなければならない。普及率100%、トラブル発生件数ゼロなどのように、意味が自明な数字を直接設定する場合を除いて、過去の実績や関連データ・情

報に基づくのが基本であろう。それをベースに関係者間の交渉によって最終的な目標値を決めるのも実績測定型評価において自然な流れである。何れにせよ、ハトリーも強調しているように、その根拠は明確にする必要がある。つまり未達によって評価が低くなることを避けるために低い目標値を設定したり、逆に予算増を求めて高い目標値を設定したりすることが常に起こりうるからである(Hatry 1999,2007)。

この問題は、個々の指標値の目標達成度合いの解釈に止まらない。目標値の性格が異なれば、目標達成度合いの指標間・プログラム間の比較ができないからである。

さて、目標管理型政策評価の事前分析表の「測定指標の選定理由及び目標値（水準・目標年度）の設定の根拠」欄あるいは評価書・実施計画において目標値の性格・設定根拠について何らかの情報が提供されている割合は、今回の調査によれば、目標値448個（達成したか否かが判断できる目標の数）のうち159個（35%）に止まっている。

行政事業レビューでは、設定された成果指標の目標値49個のうち、何らかの説明があるものは僅か8個（16%）であった。

③指標間の比較の妥当性—フロー的指標とストック的指標の区別

フロー（一定期間に発生する量）を測る指標とストック（ある時点に存在する量）を測る指標は、区別して扱う必要がある。指標値の増減が意味するところが全く異なるからである。一般に、フロー指標は行政の注力に応じて値が上下するのに対し、ストック指標は一定の注力を続ければ確実に上昇し、注力を止めても横這いとなるであろう。フロー指標の推移とストック指標の推移を無造作に比較することは避けなければならない。

実績指標には、狭義フロー・ストックの何れにも当てはまらないものも多いが、通常のフロー・ストック概念を拡張して、同じ量（質的側面も含む活動量）のプログラムを実行することが指標値の維持をもたらすのが自然であればフロー的、指標値の上昇をもたらすのが自然であればストック的とする。このようにすれば、評価で使われる指標はほぼすべてが、どちらかの属性をもつだろう。

フロー・ストックの別は、指標間・プログラム間で時系列変化や目標達成度合いを比較する際はもちろん、個々の指標の変化を読んだり、目標値を設定したりする際にも留意が必要である。ストック的指標をフロー的指標に変換する（差分を計算する）、フロー的指標をストック的指標に変換する（累計を計算する）などの方法により、すべての業績指標をフロー的な指標か、ストック的な指標のどちらかのタイプに統一できればこのような問題は回避できるが、プログラムによってフロー・ストックの何れの実績指標が適しているかは異なるであろう。

2011・2012年度の府省の実績評価・目標管理型政策評価では、目標達成状況を何段階かに評定している（プログラム間で目標達成状況を比較していることになる。目標達成状況の評定については⑥で取り上げる）7府省における取り扱いをみると、いずれの機関においてもフロー的指標とストック的指標が混在している状態であった。

④指標値の増減パターンの想定 の妥当性

実績指標の推移について「一定のペース」を想定する場合がある。例えば、中長期の目標値と中間時点の実績値を比べてプログラムの進捗が順調か否かを評価する場合、中長期の目標値に基づいて（個別の事情などを想定せずに）中間時点の目標値を定める場合、過年度から現年度までの実績値をもとに将来の目標値を設定する場合などである。

多くの場合¹⁵、当初値から目標値まで直線状の一定ペースを想定して按分しているが、このようにグラフ上で直線上に推移する、つまり等差的な軌跡を描くことを機械的に前提することが妥当であるとは限らない。一般に、等差的な推移すなわち増加率が漸減していくほかに、定率で等比的すなわち指数関数的に上昇することも一定のペースである。

どちらの増減パターンに従うと考えるのが自然なのか、評価対象プログラム・指標の性格次第であることはいうまでもない。例えば、人為的な要因に実績値が比例するようなものであれば前者であろうし、複雑な波及過程を経てアウトカムが発現するものであれば後者の考え方が自然であろう。

この認識がないまま、中間時点の評価の際に、例

えば一律直線状の軌跡を前提するとすれば、個々のプログラムの評価、その結果のプログラム間比較の両局面で危険があると言わざるを得ない。

⑤達成度計算の妥当性

数値で測定する指標を設定し、目標も数値で設定すれば、目標達成状況も数値で表すことができる。目標管理型政策評価では、標準様式に達成度を計算する欄こそないものの、2011・2012年度は4府省が、目標の達成度合いを何段階かに評定する際に数値の達成度を計算している（目標達成状況の評定は⑥で取り上げる）と考えられる。また、行政事業レビューシートには「達成度」を計算して記載する欄が設けられている（ただし今回の調査では、設定された目標値の51%しか計算されていない）。

さて、一般に実績測定型評価において目標達成度（呼称は達成率、到達率など様々である）を計算する場合、達成の度合いをどのような2つの量の比（割合）で測るかによって、2通りに大別される。両者は同じ目標達成度といっても、実は何に対する達成度合いを示しているかが異なる。

1つ目は、目標年度の実績値を目標値そのものと直接比べるもので、目標年度（期末）の目標値で実績値を除いたものである（本稿ではこれを「期末型」の目標達成度と呼ぶ）。

もう1つの方法は、基準値（期首一事業や計画の開始年度、目標値を設定した時点などにおける実績値）から目標値までの、目標達成に必要な変化量に対する達成度を計算するものである。目標としての変化量で実績の変化量を除して、「期首の基準値からどれだけ期末の目標値に近づくことができたか」を明らかにする（本稿ではこれを「期間型」の目標達成度と呼ぶ）。

シンプルな期末型の計算の特徴は、過年度の実績に全く関わりがないことである。過去1年間や数年間、あるいは目標を設定した時点からの期間の進捗・改善の度合いが反映されない。パフォーマンスが悪化したり水準が低下したりしても、その事実は達成度の数字自体には表れない。

また、期末型達成度をストック指標について計算する場合は、さらに注意が必要である。期末型達成度は定義の通り、（実績も目標も）形式上ゼ

ロに対する大きさ同士の割り算であるから、ストックがゼロであった時点から現在に至るまでの（期間型達成度に相当する）達成度でもある。すなわち、フロー指標の場合とは異なり、当該年度のパフォーマンスを直接表すものとはならない。

期間型の達成度にはこれらの問題はないが、最終的な目標値自体の達成の程度を直接知ることにはできない（達成したか否かはわかる）。したがって、過年度や途中経過よりも最終時点・現時点の状態が重要な場合（例えば、目標値が絶対に達成すべき義務的な水準の場合や、高水準の維持や低下を防ぐ意味で目標水準を定める場合など）には適さない。

自治体の実績測定型の評価では両方の型の計算が見られるが（小野 2008）、多くの場合に一律にどちらかの計算がなされていること自体が問題である。個々の指標の目標達成状況の吟味に無理があるケースがある可能性が高く、結果的に比べられないものを比べ、数えることになるからである。

府省の評価においては、行政事業レビューシートの達成度欄は一律期末型の達成度（各年度の実績値÷目標年度の目標値）を求めている。目標管理型の政策評価においては、（上述の通り非明示的に）達成度を計算している4府省のうち3機関が一律に期末型の計算をしており、1機関だけがこの2種類の達成度計算を指標ごとに使い分けている。

ただし、その1機関でも、両者の計算結果がそのまま比較可能なものとして取り扱われている（次の⑥において述べる）という問題がある。

府省において達成度計算をする場合1つの例外を除いて期末型の計算をしている背景には、自治体では総合計画の進行管理のための評価として（同じ期間に係る）期間型達成度を自然に選択する人が多いのに対し、府省ではそのような計画がない事情もある。しかし、ストック的指標でも単年度のパフォーマンスを、フローの指標でも充実・改善状況を評価している場合は当然あり、期末型達成度を一律に計算することには問題がある。もちろん目標達成度合いを数値で把握すること自体は評価すべきことであるが、丁寧な取り扱いが必要である。

なお、期末型・期間型ともに例外的な取り扱いが必要な場面（目標値がゼロの場合の期末型、現

状態維持が目標の場合の期間型など）があり、適切な対応が求められる（小野 2011b）。

⑥目標達成度の評価・集計の妥当性

個々のプログラムについて目標の達成状況を3段階や5段階に評価する（例えばA,B,Cの3段階評価など）ことは、結果をわかりやすく要約するために有効な手だてである。どのプログラムも共通の基準で評価すれば、プログラム間の比較や評価の段階別の集計ができる。

評価の際に重要なことは、評価の区分の意味が明確かつ自然であること、その評価区分がどのプログラムにおいても同じ意味をもつことの2点であろう。例えば抽象的で曖昧な判断基準のもとで各プログラムの担当者が個別に判断して集計したところで、客観的とはいえない。目標達成度の数値に基づいて評価するのが望ましいことはいうまでもない。もっとも、上述の①～⑤の妥当性が不十分であれば、当面曖昧な個別判断で評価しておく方が得策かもしれない（もちろん本来は取るべきでない消極策だが）。

目標達成度の評価は多くの自治体で行われているが（小野 2008）、府省の目標管理型政策評価・実績評価でも7機関（1府6省）が行っており（2011・2012年度）、その方法を表1に掲げた。評価を行う単位は、評価書を作成する単位である施策・施策目標・業績目標（機関によって異なる）、個々の指標、その両方という3通りがある。また、備考に示した通り、評価結果の段階毎に数を集計したり、評価結果の一覧表を作成したりしている場合がある。集計表・一覧表の作成は、個々の評価が同等かつ比較可能であることが前提されていることになる。

しかしながら、これら7機関の評価には大小様々な課題があると言わざるを得ない。何点か指摘するとすれば、定性指標の判断基準が抽象的で曖昧（機関B,C,D,E）、判断基準が異なる定量指標の評価と定性指標のそれを同一視（機関C）、定量的な判断基準しかなく定性指標での判断が不明（機関A）、達成度の数値による区分の根拠が不明確（機関A,C,E,F,G）、評価書ごとに判断基準が異なるためプログラム間の比較ができない（機関

表1 目標達成状況の評価の現状

機関	評価書（施策、施策目標、業績目標）単位	指標単位	備考
A	4段階：目標達成度の平均が100%以上、80%以上、80%未満（達成指標あり）、80%未満（達成指標なし）	—	集計表あり。数値化されてない指標・目標値あり。
B	3段階：全て目標達成、総合的に達成度合いが半分超、達成が十分でない	3段階：達成、半分以上達成、達成が十分でない	数値化されてない指標・目標値あり。
C	（定量的指標）3段階：目標達成度が80%以上、50%以上、50%未満 （定性的指標）3段階（定量的指標と共通）：共通の判断基準に基づく個別判断	—	一覧表あり。
D	5段階：共通の判断基準に基づく個別判断	—	集計表あり。
E	—	（定量的指標）3段階：目標達成度が90%以上、50%以上、50%未満 （定性的指標）3段階（定量的指標と別）：共通の判断基準のない個別判断	一覧表あり。達成度計算は期末型・期間型の使い分け。
F	3段階：A+B数のうちA数の割合が80%以上、50%以上、50%未満	2段階（+判断できない）：実績値が目標達成に向けた成果を示している(A)、示していない(B)	集計表あり。
G	—	4段階：目標達成度が100%以上、80%以上、60%以上、60%未満 など（評価書ごとに定めている模様）	

（出所）17府省の2011・12年度の目標管理型評価・実績評価（成果重視事業の評価は除く）の評価書に基づいて筆者作成。

G)、期末型達成度・期間型達成度の計算が混在したまま同じ判断基準で評定（機関E）など。

なお、達成度の数値に基づく評定の比較・集計が意味をもつためには、個々の指標の達成度が比較可能であることが前提条件である。政府では、これから「政策の特性に応じた全政府的に活用できる共通的な評価結果の表示方法」が導入されようとしているが¹⁶、その際には各府省の個々の目標達成度や目標達成状況の判断を比較・集計に耐える品質にする必要があろう。

4. 総括と今後への示唆

最後に、以上の調査結果と考察を総括し、続いてそこから得られる示唆を述べる。

（1）計量・計数に係る課題解決の必要性

政府の実績測定である目標管理型評価及び行政事業レビュー（の実績測定機能）の測定の品質には多くの課題がある。プログラムを評価するのに相応しく（妥当性を有し）、実績を測定し目標達成状況を評価できるような指標を設定する、実績と照らし合わせて達成状況を判断できるような目標値を設定する、実績値と目標値を「妥当な方法で」（本稿で提示した各局面の条件を満たすように）比較する、という各段階において、明らかな課題があると言わざるを得ない。

一方、制度導入から優に10年以上が経過した政策評価制度の実績測定において、これまで指標の品質に起因して特段の問題（品質の低い評価情報によって誤った意思決定がなされた、など）が起きていないとすれば、それは本当の意味で評価結果を反映させようとしていないから、という皮肉な事態かもしれない。Van Dooren et al. (2010) は、実績測定の品質を重視すべき理由を①測定の品質が低いことにユーザが気づけば使わなくなり、使われない測定作業は時間の無駄となり、信用回復も難しくなる、②品質の低い評価情報を使えば、間違った意思決定や行動に繋がる、という2種類に分類している。政府の実績測定型評価ではまだ、幸いにも後者の事態が起きていないのかも知れないが、評価の活用によるPDCAサイクルの構築と

いう現在の政府の方針のもと、実績測定の品質向上は急務であろう。

（2）今後の実績測定に向けて

①実績測定の制度について

現行制度において対策が必要であるほかに、今回の調査からは、制度自体の課題が示唆される。

まず、一律と多様性がともに過度になっている可能性である。定量指標・アウトカム指標を設定して目標管理を統一フォーマットにより行うという一律の原則は、目標管理型評価・行政事業レビューの何れにおいても綻びを来しているように見える。原則からの逸脱度合いが府省によって大きく違う現状は、府省の姿勢の問題もあろうが、やはり担当政策分野の違いが背景にあろう。

一方、政策評価制度では各府省が政策の特性等に応じて適切な方式を用いるという方針のもと、目標管理型政策評価の同じ基本様式を用いていても、評価対象のレベル（政策体系における階層）、基本様式の各欄の記述スタイルなど、およそ様々な状態である。評価指標の呼称や、政策体系の各層の呼称など用語も異なる。2011年度実施プログラムに関する2012年度に実施した評価を、2011年度評価・2012年度評価の何れと呼ぶかも、2派に分かれるという現状である。担当政策分野の違いを考慮に入れつつも、これらは一定程度統一すべき事項であろう。

次に、実績測定の目的の再確認の必要性である。目標管理型政策評価・行政事業レビュー（実績測定部分）の双方において、評価書やシートを文書として作成してはいても、本来の実績測定とは乖離が大きいものが少なくない。各府省が実績測定を行う意味を共有しなければ、形式を整えても仕方がない。

目標管理型政策評価と行政事業レビューの関係も、現状を踏まえた再検討が必要だろう。行政事業レビューの本旨は事業仕分けの内生化・定常化であり、実績測定とは異なる。政策評価制度との連携が求められる中、施策レベルと事業レベルの評価を繋ぐ、更には施策レベル・事業レベルを同時に管理する実績測定とするために現在の仕組みとなったが、事業レベルの評価を踏まえた施策レベルの評価に必ずしもなっていないのが現状である。

政策評価制度の実績評価方式については、政策評価制度の設計がなされた「政策評価制度の在り方に関する最終報告」において、政府にはそれまで経験がなく、知識・情報の蓄積も少ないことから目標設定や実績測定の試行を重ねる必要があり、数年を経て本格的な実施に移行することが求められていた（政策評価の手法等に関する研究会 2000）。早期に評価法が制定されたことから試行期間はほとんどなかったが、導入後10年以上が経った今、測定の品質はもちろん、実績測定の目的の再確認・再検討まで遡って改善策を講ずる必要があると思われる。

②プログラム評価の必要性

プログラム全体の整合性を問い、因果関係を踏まえて正味の効果を定量的に明らかにするような本来の（米国で見られるような）プログラム評価は、政策評価制度の設計において「総合評価方式」として想定されていたにも関わらず（政策評価の手法等に関する研究会 2000）、ほとんど実施されていないのが現状である。

本稿でも述べた通り、実績測定とプログラム評価は互いに補完する関係にある。政策評価制度においても、そのような理解に基づいて、それぞれが標準的な評価方式として用意されたはずである。プログラム評価の不在は、「測定」が主たる任務である実績測定（実績評価・目標管理型政策評価）に「評価」を求める「評価と測定の混同」（山谷 2012）の一因となっていると考えられる。実績測定への過大な期待・要請は、防御的反応や形骸化を招く誘因ともなる。各府省で一般政策を対象に本来のプログラム評価が多少とも行われるようになれば、実績測定型評価の役割もまた明確になると考えられる。

また、2013年度時点で政府の政策評価制度の当面の課題として掲げられている「結果（エビデンス）に基づく政策評価」や「（目標管理型）政策評価の重点化による質の向上」については、主として現行の目標管理型政策評価を補うものとして想定されているが、本来はこれまでほとんど行われていない、評価デザインと統計解析によるプログラム評価を試みるべきと考える。

③評価における数字の誤用・濫用

本稿で取り上げた定量指標を巡る問題点は、数字の誤用・濫用、いわゆる数字のウソの問題と捉えることができる。ここで、数字の誤用・濫用とは「数字を、意図的であれ、無知や不注意によるものであれ、正当でない、あるいは間違っている結論をもたらすような使い方をすること」（Spirer et al. 1998）である。

数字の誤用・濫用は社会統計学の分野では古くから注目されており、社会統計学の主要な問題点がほとんど出尽くした（竹内 1971）とされる19世紀末から20世紀はじめにかけてのドイツ社会統計学派の議論で既に取り上げられているが、これらの問題点は「数理統計学の立場に立ってしまうと、視界から消えうせてしまう」（内海 1979）ため、数理統計学の急速な進歩・隆盛の一方で、統計学の分野では少数の例外（Reichmann 1962、Spirer et al. 1998、上田 2003など）を除いて本格的な研究がなされず、専ら一般読者向けの読み物（最初期のもので最も広く読まれているHuff (1954) ¹⁷を始め多数）が様々に読者の注意を喚起するという状況が続いている。

一方、近年急速に取り組みが広まった国・自治体の政策評価・行政評価において、定量指標など大量の数字が使われようになる中、古くから知られている数字の誤用・濫用の問題があらためて顕在化している観がある。

そもそも政策評価・行政評価において数字や統計データが用いられるのは、有効性・効率性などを客観的に把握するためと言ってよい。ここでいう客観性とは、第一義的には個々の主観から独立であること（普遍性）だが、評価において求められる客観性にはもう1つ、特定の立場にとらわれないこと（不偏性）がある¹⁸。評価で使われる数字は、後者の客観性とも無縁ではない。政策評価・行政評価のほとんどは内部評価だからだ。評価における数字の誤用・濫用は、この何れかの意味で客観性が損なわれることを意味する。

評価において数字の誤用・濫用が発生する「過程」は、①数字の作成者・提供者が適切でない情報を伝えることを意図ないし自覚しているか否か（多くの場合、不偏性に関わるだろう）と、②結果として伝えられる数字自体に問題があるか否か

表2 数字の誤用・濫用の発生過程

意図・ 自覚	数字自体 の問題	分類
あり	あり	A：ウソの数字をつくる (架空・作為・捏造)
あり	なし	B：数字を都合よく使う (恣意的利用・隠蔽)
なし	あり	C：数字を図らずも誤って使う (非意図的過誤)
なし	なし	D：数字を独り歩きさせる (補助情報の不足・欠落)

(出所) 筆者作成。

(多くの場合、普遍性に関わるだろう) という2つの観点から、表2の4通りに分類できる。数字の誤用・濫用について統計学者が体系的に論じた上記の文献では、誤用・濫用の問題が発生する「過程」とそのような問題が発生する「場面・対象」を同じ次元で分類し体系化を試みているのであるが、筆者は（少なくとも政策評価においては）両者を明確に2つの次元として区別し、場面・対象（例えば目標管理型評価における指標設定、目標値設定、目標達成状況の比較の様々な局面）ごとに、表2の構造をもつウソの発生過程を組み合わせて考察すべきと考える。

本稿で明らかにした通り、実績測定型評価における計量・計数の現状は楽観できないものであり、問題の構造を明らかにした上で対策を講ずる必要がある。実績測定における数字のウソは、本来致命的である。数字の質の向上に向けて、更なる研究が必要である。

謝辞

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

注記

1 performance measurementについて、本稿では実績測定という訳語を用いる。業績測定（もしくは業績評価）という訳語も多く用いられる。

- 2010年度に、行政刷新会議によって事業仕分けの内生化・定常化として導入・実施された。2013年度からは行政改革推進会議の指揮下で実施される。
- 業績指標とも呼ばれる。実績測定型以外の評価も含む行政評価・政策評価分野全体では評価指標という語も広く使われる。プログラム評価において、インパクト推定のために統計解析を行うような場面では、実績指標は変数となる。
- 米国では、政府の外からのチェックとしてジョージ・メイソン大学のチームによる米国GPRAの網羅的かつ継続的な評価作業（Ellig et al. 2012）などがある。
- 2013年2月以降の経済財政諮問会議及び行政改革推進会議での議論を受け、6月14日に閣議決定された「骨太の方針」（経済財政運営と改革の基本方針）にPDCAサイクルの徹底、政策評価の重点化と行政事業レビューとの連携強化が書き込まれている。この方針のもと、2014年度からの全政府的な実行に向け、政策評価独立行政法人評価委員会の政策評価分科会の審議や各府省の意見を踏まえながら総務省行政評価局を中心に検討が進められる予定である（2013年6月21日開催の政策評価分科会での議論）。
- 各機関が公表時に評価書を並べている順序に従い、最初に挙がっているものを1つ目の標本とし、残り9個の標本をできるだけ等間隔になるように抽出した。行政事業レビューシートについても同様。各機関ごとに（評価対象プログラムが10未満の機関を除いて）同数の標本を抽出しているため、全体の集計は、各機関に同じ重みを与えた政府の全体像に相当する。
- 行政改革推進会議においても、定量化・指標デザインの改善が必要であることが議員から指摘されている（田中 2013）。
- 竹内編（1989）の「データの精度と信頼性」の項参照。
- Wholey（2007）は適時性（timeliness）、Rossi et al.（2004）は感度（sensitivity）を妥当性・信頼性と並ぶ条件として挙げている。適時性は、指標の設定・測定の何れの場面に係るかにより、本文で後述する論理的妥当性・操作的妥当性の何れかに相当する。感度とは「測定されている事象に変化や差異があるときに、対応してその測定尺度の値が変化する程度」のことであり、本稿でいう操作的妥当性に相当する。
- プログラム評価の分野では、インパクトを定量的に

明らかにするという観点から、(シンプルな指標による測定を旨とする)実績測定の分野とは若干異なる流儀で妥当性・信頼性に関する整理を行う。プログラム評価のための評価デザインと統計解析法を統合して体系的に述べた数少ないテキストであるLangbein (2012)は、妥当性・信頼性を①internal validity、②external validity、③measurement validity & reliability、④statistical validityという4つに分類している。実績測定における妥当性・信頼性と対応関係を示せば、本稿でいう論理的妥当性は上の①に、操作的妥当性は③の一部(validity)に、信頼性は②と③の一部(reliability)に概ね相当する。上の④は変数間の関係に関する妥当性であり、通常の実績測定には該当しない。

- 11 今回の調査の対象とした2012年度に行われた評価では、評価対象の2011年度実施プログラムの事前分析表は8機関しか用意されていないが、5機関は評価書や評価実施計画書に該当情報を記載する欄等を設けている。
- 12 これら諸点についての詳細は小野(2011b)を参照されたい。また、小野(2008)はこれら諸点について都道府県の実績測定型評価における現状について述べている。
- 13 「過年度の水準より向上」、「現状維持」という目標であっても達成度合いを精密に判定する有意義な試みとして、国立国会図書館の2011年度活動実績評価を挙げる。
- 14 例えばリドレー及びサイモンなど米国行政学の動向を踏まえて「能率」(今日の評価論でいう「効率」に相当)を論じた西尾(1976)は、政策の目標値には期待値・充足値・限界値(本稿に述べた目標値の性格3種類にほぼ相当)があると指摘している。
- 15 筆者が近年目にした、府省と自治体の実績測定型評価の該当するケース(目標年度の目標値まで「一定のペース」を具体的に想定するケースなど)では、1つを除いた全て。
- 16 2013年6月21日開催の政策評価分科会での議論。
- 17 Huff(1954)は世の中の「統計のウソ」を詳細に論じた最初期の著作であるだけでなく、歴史上最も多く読まれた統計学書である(Steele 2005)。刊行後50年を経た2005年に米国の統計学会誌Statistical Scienceが特集を組むなど、統計学者もその意義・功績を認めている。

18 2種類の客観性については竹内(1971)の整理を参考にしている。

参考文献

- 上田尚一(2003)『統計の誤用・活用』、朝倉書店
- 内海庫一郎(1979)「フランクフルト学派統計学の略図」、北海道大学『経済学研究』、29(1):1-23
- 小野達也(2003)「社会指標による政策効果測定に関する一考察—時系列指標を比較する方法」、『日本評価研究』、3(2):169-178
- 小野達也(2008)「都道府県の評価の10年—定量評価のこれまでとこれから」、『日本評価研究』、8(1):19-38
- 小野達也(2011a)「自治体ベンチマーキングと指標の比較可能性」、『日本評価研究』、11(2):13-30
- 小野達也(2011b)「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)
- 政策評価の手法等に関する研究会(2000)『政策評価制度の在り方に関する最終報告』
- 総務省(2011)「目標管理型の政策評価の改善方策に係る試行的取組について」、『平成23年度における政策評価の実施について』総評政第14号行政評価局長通知別紙
- 竹内啓(1971)『社会科学における数と量』、東京大学出版会
- 竹内啓編(1989)『統計学辞典』、東洋経済新報社
- 田中弥生(2013)『行政事業レビュー改善に向けて—より効果的・効率的な仕組みへの提案』、行政改革推進会議2013年度第2回資料
- 田辺国昭(2001)「実績評価の課題」、『都市問題研究』53(10)
- 田辺国昭(2005)「中央省庁における政策評価の現状と課題—期待と失望のスパイラルを超えて」、『政策形成支援のための政策評価—NIRA型政策評価モデルの提言』、総合研究開発機構
- 田辺国昭(2006)「政策評価制度の運用実態とその影響」、『レヴァイアサン』38
- 田辺国昭(2007)「政策評価の現状と課題—評価を有効にする組織条件」、『評価クォーターリー』1
- 西尾勝(1976)「効率と能率」辻清明他編『行政学講

- 座』3、東京大学出版会
- 古川俊一（2002）「公共部門における評価の理論・類型・制度」、『公共政策研究』2
- 山谷清志（2006）『政策評価の実践とその課題—アカウンタビリティのジレンマ』、萌書房
- 山谷清志（2012）『政策評価』、ミネルヴァ書房
- De Bruijn, H. (2007). *Managing Performance in the Public Sector 2nd Edition*. London: Routledge
- Ellig, J., McTigue, M. and Wray, H. (2012). *Government Performance and Results: An Evaluation of GPRA's First Decade*. New York: CRC Press.
- Greiner, J.M. (1996). Positioning Performance Measurement for the Twenty-First Century. In Halachmi, A. and Bouckaert, G. (Eds) *Organizational Performance and Measurement in the Public Sector: Toward Service, Effort and Accomplishment Reporting*. London: Quorum Books.
- 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. (2007). *Performance Measurement: Getting Results 2nd Edition*. Washington, D.C.: Urban Institute.
- Hatry, H.P. (2013). Sorting the Relationships Among Performance Measurement, Program Evaluation, and Performance Management. *New Directions for Evaluation*, 137.
- Huff, D. (1954). *How to Lie with Statistics*. New York: W.W.Norton & Company (高木秀玄訳 (1968)『統計でウソをつく法—数式を使わない統計学入門』、講談社)
- Langbein, L. (2012). *Public Program Evaluation: A Statistical Guide 2nd Edition*. New York: M.E.Sharpe.
- Newcomer, K.E. (1997). Using Performance Measurement to Improve Programs. *New Directions for Evaluation*, 75.
- Nielsen, S.B. and Hunter, D.E.K. (2013). Challenges to and Forms of Complementarity Between Performance Management and Evaluation. *New Directions for Evaluation*, 137.
- Osborn, D. and Gaebler, T (1992). *Reinventing Government: How the Entrepreneurial Spirit Is Transforming the Public Sector*, New York: Penguin Books (日本能率協会訳 (1995)『行政革命』日本能率協会マネジメントセンター)
- Poister, T.H. (2004). Performance Monitoring. In Wholey, J.S., Hatry, H.P. and Newcomer, K.E. (Eds). *Handbook of Practical Program Evaluation 2nd Edition*. San Francisco, Jossey-Bass.
- Reichmann, W.J. (1962). 浜崎啓治・是永純弘訳 (1967)『統計—利用と誤用』法政大学出版社
- Rossi, P.H., Freeman, H.E. and Lipsey, M.W. (2004). *Evaluation: A Systematic Approach 7th Edition*, Sage Publications. (大島・平岡・森・元永監訳 (2005)『プログラム評価の理論と技法』、日本評論社)
- Spirer, H.F., Spier, L. and Jaffe, A.J. (1998). *Misused Statistics: 2nd Edition*. New York: Marcel Dekker, Inc.
- Steele, J.M. (2005). Darrell Huff and Fifty Years of How to Lie with Statistics. *Statistical Science*, 20 (3)
- Swiss, J.E. (1995). Performance Monitoring Systems. In Ammons, D.N. (Ed). *Accountability for Performance: Measurement and Monitoring in Local Government*. Washington, D.C.: International City/County Management Association.
- U.S. General Accounting Office (1992). *Program Performance Measures: Federal Agency Collection and Use of Performance Data*. Report GAO/GGD-92-65. Washington, D.C.: U.S. General Accounting Office
- Van Dooren, W., Bouckaert, G. and Halligan, J. (2010). *Performance Management in the Public Sector*. London: Routledge
- Wholey, J.S. (2007). Quality Control: Assessing the Accuracy and Usefulness of Performance Measurement Systems. In H.P.Hatry, *Performance Measurement: Getting Results 2nd Edition*. Washington, D.C.: Urban Institute.

(2013.9.19受理)

Performance Measurement in the Evaluation Practices of Japanese Government ～Exploring the Quality of Measurement～

Tatsuya Ono

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

Abstract

Performance measurement has been widely used in the “Policy Evaluation System” of Japanese government. Since 2011, this system has been linked to “Administrative Operations Review”, another system of performance measurement newly implemented. Although Japanese Government intends to use its performance measurement systems just as Western countries do, there are some difficulties. This article classifies the problems of performance measurement practices in Japanese government and explores the way of acquiring the quality of measurement. This article also discusses the forms of complementarity between performance measurement and program evaluation, that is a weak point in the evaluation systems of Japan.

Keywords

performance measurement, performance measurement in the “policy evaluation system”,
target setting and benchmarking, administrative operations review

【研究論文】

政策評価制度における評価の利用： 「評価の利用」概念の解明と現状の分析

田中 啓

静岡文化芸術大学

hiraki@suac.ac.jp

要 約

国の行政機関が政策評価に取り組み始めてから11年以上が経過したが、近年、評価が十分に利用されていないことが問題視されている。「評価の利用」はその定義や概念が明確にされていないため、評価の利用の実態も十分に解明されるに至っていない。そこで本稿では、「評価の利用」概念を解明し、評価の利用を実証的に分析するための枠組みを構築することを主眼とした。参考にしたカークハートのモデルは、プログラム評価の影響が波及する経路を概念的に理解するための社会的視点のモデルである。本稿では、これに修正を加え、国の政策評価の利用を分析するための組織的視点のモデル（修正KHモデル）を構築した。このモデルを用いて国の政策評価の利用状況について試論的な分析をおこなったところ、評価の利用者は施策担当者と管理職にほぼ限られていることや、評価の用途も施策理解が主であり、意思決定等に関して評価が積極的に利用されていないこと等が示唆された。今回構築した理論モデルは、評価の利用を実証的に分析するために有効な枠組みであり、このモデルを用いた詳細な実態分析が必要である。

キーワード

政策評価制度、政策評価法、評価の利用、実用的評価

1. はじめに

国の行政機関が評価に取り組み始めて久しい。起点をどこに置くかにもよるが、「行政機関が行う政策の評価に関する法律」（以下、政策評価法）が施行された2002（平成14）年4月を基準にすれば、少なくとも11年以上が経過したことになる。

11年を振り返ると、政策評価制度を取り巻く状況には、特筆すべき変化や出来事もあった（近年の例としては、2009年の政権交代とその後の事業仕分けの実施、2011年の東日本大震災の発生な

ど）。この間に政策評価制度の「見直し」が実施されたり、評価の実施を義務づける対象が拡大されたりするなど、政策評価制度の規定や運用には一定の変更が加えられてきた。しかし、政策評価制度の基本的な設計思想や構造自体には大きな変化もなく、現在に至っている。ただしそれは、政策評価制度に問題がないことを意味しているわけではない。むしろ近年は、国の政策評価について問題点や課題がさまざまな形で指摘されるようになっている。

政策評価についてしばしば指摘されるのは、評

価が十分に利用されていないという問題点である。もしも評価があまり利用されていないとすれば、評価の実施に関わる数多くの行政職員の労力（ひいては国民の税金）が浪費されていることになる。だとすれば、「政策評価が利用されていない」と指摘されていることに対して、行政関係者は無関心であるべきではない。評価の利用実態を把握し、もしもそれに問題があれば、その改善に取り組むことが必要である。

ところで、評価の利用について検討する際に問題となるのは、「評価の利用」とは一体どのような状態（または行為）なのかということである。私たちはごく普通に「評価の利用」という表現を使用しており、通常はそれで大きな問題はない。しかし、評価を利用しているのかいないのかを真剣に検討しはじめた途端、「評価の利用」が実に幅広い解釈を許容する表現であることに気づかされる。

例えば、「評価を利用する」とは、評価の「何」を利用することなのだろうか。評価の「結果」だけなのか、あるいは評価の「過程」も含むのか。そもそも評価の「結果」とは、評価書に記述されたことを指すのだろうか。だとすれば、評価を実施する過程で得た知見は評価「結果」ではないのか。さらに「利用」についても、具体的にどのような行為を意味するのだろうか。評価結果を利用して何らかの判断をおこなうことが「利用」なのだろうか。だとすれば、評価結果を見て「参考にする」のみに留めた場合、「利用」とは言えないのだろうか。

このように、「評価の利用」とは、さまざまな解釈が成り立つ概念であるにもかかわらず、その意味は、これまであまり厳密には追究されてこなかった。恐らく従来は、評価を導入・実施することに主眼が置かれており、評価手法の開発や評価の実施方法の理解により強い関心が寄せられていたからであろう。だが、明らかに行政機関における評価の潮目は変わっており、少なくとも実務面の関心は、評価の実施から利用へと移行しつつある。

本稿の究極的な関心の対象は、国の政策評価の利用実態（の把握）にある。ただし、以上で述べた事情から、「評価の利用」概念を明確にすることを避けて、そこに到達することはできない。そ

こで本稿では、まずは「評価の利用」を概念的に解明し、評価の利用実態を把握・分析するための理論的枠組みを構築することを主眼としたい。その上で、構築した理論的枠組みに基づき、国の政策評価制度における評価の利用実態を試論的に検討することにする。その意味では、本稿の目的は、政策評価における利用実態を明らかにすることではなく、そのための準備として、理論的な整理や枠組みの構築をおこなうことであることをあらかじめ断わっておく。

本稿の構成は次のとおりである。2.においては、「評価の利用」を構成する主要な要素を個別に検討し、「評価の利用」概念を理論的に解明する。その結果を踏まえて3.では、「評価の利用」を実証的に把握・分析するための理論モデルを検討する。具体的には、Kirkhart（2000）が提示したモデル（Integrated Theory of Influence）を参考にして、このモデルに所要の修正をほどこすことにより、本稿の目的に適したモデルを導入することにする。

最後に付論では、そのモデルを利用して、国の政策評価における利用の実態を定性的に検討する。ただし、本稿の執筆にあたり、国の行政機関における政策評価の実施状況について詳細な調査をおこなったわけではないことから、ここで示すのは、筆者の知見やこれまでの観察に基づく暫定的な判断の結果であることをあらかじめ断っておきたい。

なお、「利用」に類似する概念に「活用」がある。一般的に「利用」と「活用」はほとんど同意に用いられており、政策評価の領域においても両者はあまり明確に区別されていない。本稿では敢えてその慣例に一石を投じて、両者を明確に区別することにする。両者の違いは後述するが、本稿では「利用」という表現を主に使うことにしたい（その理由も後述する）。

2. 「評価の利用」の理論的整理

(1) 評価の「利用」の重要性

評価研究者の間では、評価の利用を重視すべきかどうかについて、長年にわたり論争が続けられ

てきた（佐々木 2010）。評価の利用を重視すべきかどうかは、実用的評価と科学的評価のどちらを重視するかという立場の違いに帰着する。本稿は、言うまでもなく前者の立場（評価の利用を重視する立場）に立つものである。ただし、公正を期するために、科学的評価を重視する考えも依然として根強く存在することを指摘しておく。

一方、国の政策評価制度に限って考えてみると、行政機関が政策評価を実施し、その経過や結果を議会や国民に対して示すことで、政策評価は十分にその役割を果たしているという立場もありうる。これは、政策評価によって行政機関がアカウンタビリティを果たすことを重視する立場といえよう。ただし、以下に示す政策評価法の条文からの引用をみれば、政策評価制度はそのような目的（アカウンタビリティを果たすこと）のためだけに導入されたのではないことは明らかである。

第1条「[···] 政策の評価の客観的かつ厳格な実施を推進しその結果の政策への適切な反映を図るとともに、政策の評価に関する情報を公表し、もって効果的かつ効率的な行政の推進に資するとともに、政府の有するその諸活動について国民に説明する責務が全うされるようにする」

第11条「行政機関の長は、少なくとも年1回、当該行政機関における政策評価の結果の政策への反映状況について、総務大臣に通知するとともに、公表しなければならない。」（以上、傍点は筆者）

これらの条文から明らかなように、政策評価に関する情報を公表することだけでなく、評価結果を政策に反映することもまた政策評価制度の重要な目的とされているのである。

（2）政策評価制度における「利用」概念

「評価の利用」概念を詳しく検討する前に、国の政策評価において「利用」をどのようにとらえているかを確認しておきたい。政策評価制度の公式用語においては、「政策への反映」が「利用」に相当する概念である。「政策への反映」とは、先にみた政策評価法の第1条等に表示されている表

現である。

「政策への反映」が具体的にどのような状態を意味するかは、政策評価法をはじめとして政府の公式文書には示されていない。一般的な用語法から推察すれば、政策に対して何らかの改善をほどこすことが「政策への反映」の意味するところであると考えることができる。また政策評価法には「政策の評価 [···] の結果の政策への適切な反映···」（傍点は筆者）と記述されているので、政策に反映する対象は、評価「結果」であると想定されている。

政策評価法にこのような規定があるため、総務省は毎年「政策評価等の実施状況及びこれらの結果の政策への反映状況に関する報告」（以下、「反映状況報告」）という文書を作成し、国会に対して報告している。残念なことに、総務省がこの文書において「政策への反映」をどのように定義しているかは明らかではない。ただし、この文書中には各府省における評価結果の反映状況をまとめた資料があり、その内容から判断すると、評価結果を利用して評価対象の政策を改善・見直し・廃止・休止・中止した場合だけでなく、「これまでの取組を引き続き推進」している場合も「反映件数」に含まれている。

ここで疑問が生じるのは、次の2点である。

第1に、「評価結果の政策への反映」における「評価結果」とは、具体的に何を指しているのかということである。もちろん、評価者が評価を実施し、その結果が評価書に示されるような内容としてまとめられた場合、それはまぎれもなく「評価結果」である。意思決定者は「評価結果」が示された評価書を見て、政策の改善や見直しなどの判断を下す。

ところが、評価の実施に並行して次年度の政策を立案しているような場合、評価の実施途中（すなわち評価の結論が出る前）で担当者が政策の問題点に気づき、政策を修正したとすれば、「評価結果の反映」に含まれるのであろうか。この場合、評価対象の政策は、評価の実施途中で得られた知見に基づき修正されるため、最終的な評価書には修正前の政策の問題点を指摘する必要がなくなる。文字どおり「評価結果」を利用して政策を修正したのでないとしても、評価を実施することが

政策の修正につながったとすれば、このような事例も「政策への反映」に含めるべきという考え方もありうる。

実際には、このようなケースも「政策への反映」に含めているのかもしれないが、各府省がどのような基準で「政策への反映」を把握しているかが明示されていないため、評価の「何」を政策に反映しているのかが今一つ判然としないのである。また府省間で「政策への反映」のとらえ方に違いがあるのかもどうかも不明である。

2つ目の疑問は、「反映」とは何を意味しているのかということである。先にみたように、政策の改善・見直し・廃止・休止・中止を決定した場合は「反映」に含まれる。ただしこの場合、評価結果とこれらの判断はどのように結びつけられているのだろうか。例えば、評価結果が政策の廃止や中止を明確に示唆していないにもかかわらず、その政策を廃止ないし中止した場合は、評価結果の「反映」に含まれるのであろうか。

一方、総務省による「反映状況報告」では、「これまでの取組を引き続き推進」する場合、言い換えれば、政策に特段の変更を加えない場合でも「反映」に含むこととされている。もしもそうであれば、政策評価の結果、特に問題が発見されなかったため政策に変更を加えなかった場合は、すべて「反映」といえるのだろうか。

要するに、政策についてどのような判断（改善・見直し・廃止・休止・中止・変更なし）を下すにせよ、評価結果の示した結論に沿うものであれば「反映」と言えるのであろうか。この点についても、総務省をはじめとして国が公表する資料は、何も教えてくれない。

このように、政策評価制度における「評価の利用」は明確に定義されておらず、総務省が公表している「評価結果の反映状況」も、踏み込んで考えてみると、その意味は不明である。以上を総合すれば、政策評価制度においては、「利用」概念が明確に定義されておらず、そのために政策評価の「利用」実態も明らかにされているとは言えないのが現状である。

(3) 「評価の利用」とは何か

そこで、「評価の利用」概念を独自に定義する

必要がある。そのために、「評価の利用」をいくつかの要素に分解して検討する。具体的には、①用途・目的、②利用対象、③利用者の3つの要素である。なお、この整理はWeiss (1998) を参考にしている。

①用途・目的

用途・目的とは、「評価を何のために使うか」ということである。Weiss (1998) によれば、用途・目的として「事業の理解」「意思決定」「説得・支持獲得」「アカウントビリティ」「他機関への影響」の5種類がある（図1）。

このうち利用の目的として最も一般的なのは「意思決定」であろう。これには、事業の採否や改善の決定など、何らかの判断を伴う利用形態が全て含まれる。また日本では一般的ではないが、意思決定者や資金提供者などに対して説得したり支持を獲得したりするために評価を利用することも用途のひとつとして挙げられている。ワイスは、「意思決定」と「説得・支持獲得」という2つの用途を評価の「道具的利用 (instrumental use)」としており、その他の3つの用途（「道具的利用」に対比して「概念的利用 (conceptual use)」と表現されている）と区別している。

また5つの用途・目的のうち「事業の理解」「意思決定」「説得・支持獲得」「アカウントビリティ」は、利用者が意図的に評価を利用する場合（「意図した利用 (intended use)」）であり、「アカウントビリティ」や「他機関への影響」は利用者が「意図しない利用 (unintended use)」と考えることができる。ただし「アカウントビリティ」については、それを果たすために評価を「意図して利用」する場合と、評価を実施することが結果的にアカウントビリティを果たすことにつながるという「意図しない利用」の両方の可能性がある点に留意が必要である。

②利用対象

次に利用対象とは、評価の「何」を使うのかということである。政策評価制度においては、政策評価の「結果」を利用することが想定されている。これを「結果の利用 (results use)」と表現することにする。

評価の利用を実証的に分析するためには、「結果」の内容をさらに定義しておく必要がある。ここでは、評価結果の内容を事業の「基本情報」・「実施状況」・「実施結果」と「判断・提言」の4つに分類する。

「基本情報」とは、評価を実施する際に確認（または設定）する事業の目的、内容、事業費等の情報である。評価の準備作業に近い内容であるが、評価を実施する際に、事業の内容を確認したり事業の目的を設定したりするのは欠かすことのできない行為である。その結果は、ワイスが評価の用途・目的のひとつとして挙げている「事業の理解」の対象になる。

「実施状況」とは、事業の実施や進捗の状況を定量的・定性的に把握することにより得られるあらゆる情報である。

「実施結果」は、事業を一定期間実施した後または事業の終了時に、事業の成果や実施に伴う変化を定量的・定性的に把握したものである。

最後に「判断・提言」とは、「実施状況」や「実施結果」に関する知見を踏まえて示された判断や提言のことである。「判断・提言」には評価者の価値判断が含まれている。

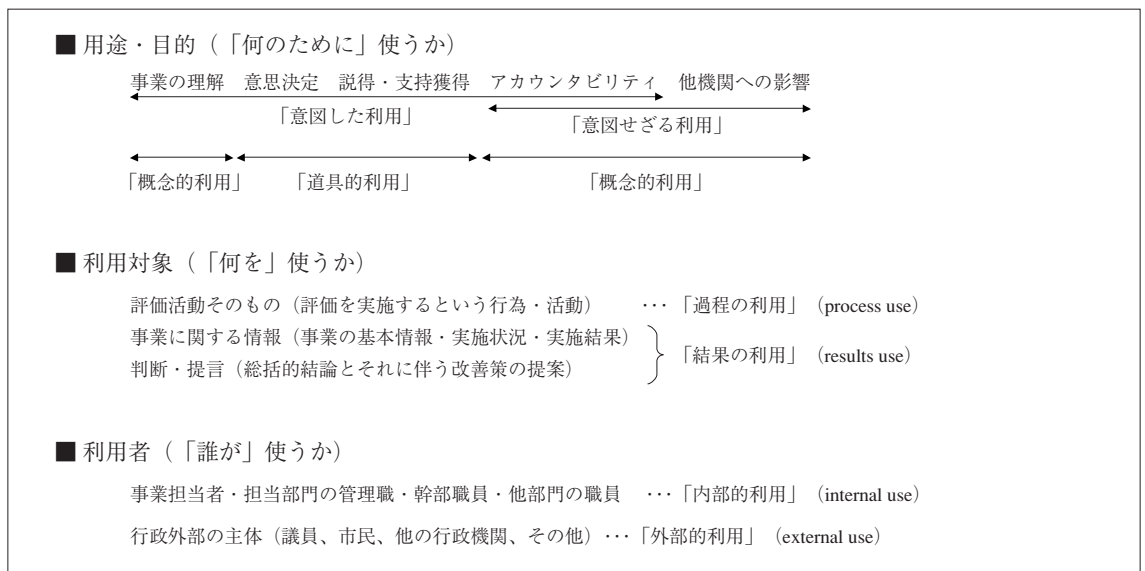
評価論に基づけば、利用対象は評価結果だけとは限らない。評価の「過程」も利用対象となりうる。評価の「過程」を利用することを「過程の利用（process use）」と呼ぶ。評価を実施するという行為・活動自体に利用価値を認める場合である。例えば、エンパワメント評価（empowerment evaluation）においては、政策の対象者や一般市民に評価に参加してもらい、これらの人びとが評価活動に参加することで、参加者自身の意識変化や学びを期待する。このような場合、利用の対象は評価を実施した結果ではなく、評価活動そのものである。これが「過程の利用」の典型例である。

③利用者

最後に利用者とは、評価を「誰が」利用するかという点である。行政内部の主体が利用する場合と行政外部の主体が利用する場合に大別される。前者を「内部的利用（internal use）」、後者を「外部的利用（external use）」と呼ぶ。

政策評価制度をはじめとして、国内の行政機関がおこなう評価に関しては、基本的には「内部的利用」が中心である。行政外部の主体（例えば、議員、市民、他の行政機関など）が利用する場合

図1 「評価の利用」の諸側面



もないわけではないが、局所的または非継続的なものに留まっているというのが実態である¹⁾。

「内部的利用」の利用者としては、「事業担当者」「事業担当部門の管理職」「幹部職員」「他部門の職員」に大別することができる。

以上で整理したように、「評価の利用」とは、少なくとも「用途・目的」「利用対象」「利用者」の3つの側面において、さまざまな要素から構成される多面的な行為・活動である。このように整理すれば、「評価の利用」を定義することの困難さが理解されよう。

「評価の利用」を簡潔な表現で定義しようとするれば、いかなる側面や要素も排除しないような抽象的な記述にならざるを得ない。教科書上の記述であればそれで良いのかもしれないが、実務や実証的研究に役立てるためには、目的や用途に応じて、その都度「評価の利用」を具体的に定義する必要がある。

(4) 「利用」と「活用」

ここまでは「評価の利用」について概念的な整理をおこなってきた。1.で指摘したように、「利用」と似た概念に「活用」がある。通常は「利用」と「活用」は明確に区別されておらず、ほぼ同じ意味合いで使用されている。だが、「利用」と「活用」を明確に区別することにより、「評価の利用」の意味を一層明確にすることができる。

ちなみに辞書（広辞苑第五版）で「利用」は「利益になるように物を用いること。役に立つように用いること。」「活用」は「活かして用いること。効果のあるように利用すること。」と説明されており、両者の違いは明確ではない。また英語で表現すれば、「利用」はuseまたはutilizationであるのに対し、「活用」はutilizationのほか、applicationやexploitationという訳語も当てはまる。

本稿で提示する解釈は、「利用」をutilization、「活用」をexploitationとする訳語に近い。本稿では、「利用」を「想定した用途・目的に供すること」、「活用」を「想定した以外の用途・目的に供すること」と定義することにする。

このように定義すれば、「評価の利用」とは、評価を実施する際にあらかじめ想定していた目的（例えば、評価結果に基づく事業の採否決定等）

に用いることであり、「評価の活用」とは、あらかじめ想定していなかった用途（例えば、評価結果を使った国民との対話等）に用いることである。両者の違いは、所期の使い道か、そうでないかという点にある。

なお、評価の用途・目的には「意図した利用」と「意図しない利用」の2種類の区分があることを既に示した。「意図した／意図しない」と「想定した／想定しない」はよく似ているが、当初想定していなかった用途に使う場合でも、意図的に用いるのであれば、「意図した利用」とみなすことができる。

本稿では、主に評価の「利用」という表現を用いることにしている。それは、あらかじめ用途を想定した上で評価を実施し、実際にその過程や結果をその用途に使用することが重要だと考えるからである。もちろん、評価の「活用」と表現されるような、評価の派生的な使用方法を否定するわけではない。ただし、現在の政策評価の問題点は、「想定した用途」にすら評価が使われていない点（あるいは、あらかじめ用途が「想定」されていない点）にあると筆者は考えており、そうであるとすれば、評価の「活用」以前の問題といえる。

こうした認識に基づいていることから、本稿では評価の「利用」という面を「活用」よりも重視することにする。以下でも、基本的には評価の「利用」面に焦点を当てて検討を進めていくことにしたい。

3. 「評価の利用」の理論モデル

(1) 利用のモデル化

「評価の利用」概念を整理することができたので、次の課題は、その概念を用いていかにして政策評価制度の実態を分析するかということである。そのために、評価の利用に関する理論モデルを導入する。ここではカークハートが考案したモデル（Kirkhart 2000）を参考にして、これを本稿の目的に使用するために修正をほどこすことにしたい。

まずカークハートのモデルを紹介する。カークハートのモデル（Integrated Theory of Influence）

は、プログラム評価を主に念頭に置いたモデルである（ただし、業績測定型の評価も対象から排除されていない）。政策評価と同様、プログラム評価においても、評価が十分に利用されていないという問題意識が専門家の間で共有されてきた。そのような問題意識を背景として、プログラム評価の利用がもたらす影響を概念的に理解するためにカークハートが考案したモデルである。

本稿がカークハートのモデルを利用する理由は次の2点である。

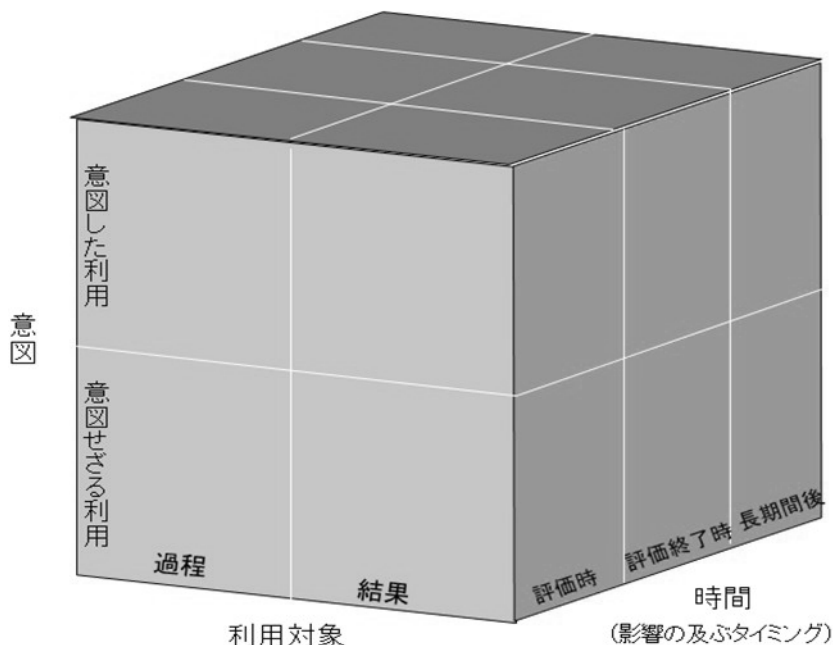
第1に、カークハートのモデルは、評価の利用が複雑で重層的な過程であることを踏まえた上で、これを解明するための理論的な枠組みを提供してくれるからである。評価の利用に関する先行研究は、利用実態を実証的に解明しようとするものが大半を占めており（例えばde Lancer Julnes & Holzer 2001など）、理論的な論考の場合には利用の定義や意義を論じるに留まるものが多く、本稿が必要とするような評価の利用の構造を分析するための理論的枠組みがほとんど存在しない。

第2に、カークハートのモデルの包括性である。評価の利用に関しては、専門家の間でさまざまな見方や教義が存在するが、カークハートのモデルは、評価の利用に関する先行研究を踏まえて、評価の利用をできるだけ包括的に分析するために考案されたものである。評価に関する特定の立場に依存したモデルではないため、評価の利用をさまざまな側面から解明しようとする本稿の目的に合致している。

カークハート・モデルでは、利用対象（source）、意図（intention）、時間（time）という3次元を用いて、評価の利用がもたらす影響の波及をとらえようとする。このモデルを視覚的に示したのが図2である。利用対象は「評価の過程」と「評価の結果」、意図は「意図した利用」と「意図せざる利用」のそれぞれ2つずつに分けられている。評価を利用した影響は、時間の経過に伴って波及していくが、このモデルでは評価時・評価終了時・長期間後の3時点に分けている。

カークハートは、評価の利用がもたらす影響が

図2 カークハートのモデル（Integrated Theory of Influence）



(出所) Kirkhart (2000)

複雑なプロセスを経て波及することを強調した上で、このようなモデルを用いることにより、評価に関係する主体や研究者が、評価がもたらす影響について、より明確に理解することができるようになる」と述べている（Kirkhart 2000）。

先述したように、カークハート・モデルはプログラム評価を主に想定したモデルなので、評価に関する専門的知識や技能を持つ主体が評価をおこなうことが、暗黙のうちに想定されている（評価の企画者や依頼者も行政府とは限らない）。またこのモデルでは、評価の利用者は明示的に示されておらず、専門家が実施した評価の過程または結果が、意図的あるいは非意図的に利用され、社会に影響を及ぼしていく波及経路に関心が置かれている。言い換えれば、評価が行政府だけでなく、社会に広く影響を与える過程を概念的に理解することを主眼とした社会的な視点のモデルである。

本稿では、このモデルを政策評価の分析に応用するためにいくつかの修正をおこなうことにする。まずカークハート・モデルでは、「利用対象」「意図」「時間」の3次元であったのを、同じ3次元であるが、「利用対象」「用途」「利用者」に変更する。

カークハート・モデルとの大きな相違点は、「利用者」という次元を導入し、代わりに「時間」を捨象したことである。カークハート・モデルでは、評価結果が時間を経て社会で利用されていく過程を分析するために「時間」の要素が強調されているのに対し、本稿では評価を実施する行政機関内で評価が利用される過程に注目するため、「時間」の要素よりもむしろ誰が「利用者」であるかが重要になる（行政機関内での利用は、ほとんどが数か月から1年以内で完結するため、時間

の要素の重要性は相対的に低い）。またカークハート・モデルでは、「意図」という次元を用いて利用が意図的であるかどうかに注目しているのに対し、本稿では、用いている定義上、利用が意図的であることは前提であり、むしろ「何のために」評価を利用するかにより強い関心を置いている。そこで、意図の代わりに「用途」という次元を導入することにした。

カークハート・モデルは社会への影響の波及経路に関心を置く社会的な視点のモデルであったが、これを修正したモデル（「修正KHモデル」と呼ぶことにする）は、行政機関が内部で実施した評価が、主に行政機関内部で利用される過程を分析するための組織的な視点のモデルといえる。

その意味では、本稿のモデルはカークハート・モデルを政策評価の分析に直接応用しようとしたものではなく、カークハート・モデルの理論的枠組みを参考にして導出した独自のモデルと説明する方がより正確であろう。

修正KHモデルで用いる3次元は、2（3）および図1で示した概念整理の内容を踏まえて、カークハート・モデル以上に細かい要素に分類した。

各次元の要素は表1に示すとおりである。「利用者」としては、施策担当者、管理職、幹部職員、他部門の職員の4つの要素に分けて考える。「利用対象」は、評価活動そのもの、施策の基本情報、施策の実施状況、施策の実施結果、判断・提言の5つに分ける。最後に「用途」は、施策の理解、意思決定、説得・支持獲得、アカウントビリティの4つの要素とする。

図1では、利用者と用途・目的のいずれの次元も5つずつの要素に分けていたが、本稿では、行政機関内部における評価の利用が関心の対象であ

表1 修正KHモデルにおける次元と要素

3 次元	要 素
a. 利用者	施策担当者、管理職、幹部職員、他部門の職員
b. 利用対象	評価活動そのもの、施策の基本情報、実施状況、実施結果、判断・提言
c. 用途	施策の理解、意思決定、説得・支持獲得、アカウントビリティ

（出所）筆者作成

ることから、利用者のうちの「行政外部の主体」と用途・目的のうちの「他機関への影響」という要素をそれぞれモデルから省くことにした。

カークハート・モデルにならい、修正KHモデルを3次元で視覚的に示したのが図3である。カークハート・モデルでは、評価が社会に影響を及ぼす経路について「利用対象→意図→時間」という経路が想定されるのに対し、修正KHモデルでは、行政機関内部において実施された評価が「利用対象→利用者→用途」という経路で効果を発揮していくことが想定される。

(2) モデルを利用した分析の枠組み

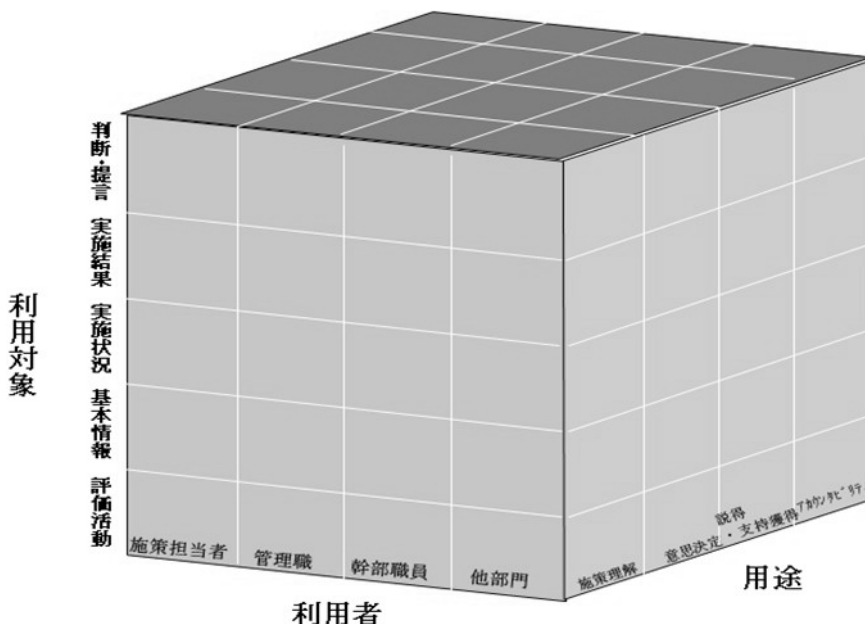
次に、修正KHモデルを利用して、政策評価制度における評価の利用実態を分析するための枠組みを設定する。修正KHモデルは、利用者・利用対象・用途の3次元からなるモデルであった。しかもそれぞれの次元は4ないし5つの要素に分かれている。したがって、3次元の要素別の組み合わせは $4 \times 5 \times 4$ の80とおりになる。この全ての組み合わせをひとつの図や表で示すことは困難なの

で、2次元ずつの組み合わせに分解して検討することにした。つまり、「利用者」×「利用対象」、「利用者」×「用途」、「用途」×「利用対象」の3とおりの組み合わせが分析の対象となる。

表2は、この3とおりの組み合わせを示したものである。Aは、「利用者」と「利用対象」という2次元の構成要素をマトリックスで示したものである。「利用者」の構成要素としては、施策担当者、管理職、幹部職員、他部門の4つである。一方、「利用対象」としては、評価活動、基本情報、実施状況、実施結果、判断・提言という5つの要素がある。Aの表では、利用者の別に評価のどのような対象を利用しているかを把握することができる。

次にBは、「利用者」と「用途」についてのマトリックスである。「用途」の構成要素は、施策の理解、意思決定、説得・支持獲得、アカウンタビリティの4要素である。この表では、各利用者がどのような用途で評価を利用しているかを把握することができる。最後にCは、「用途」と「利用対象」についてのマトリックスである。この表からは、用途別に評価のどのような部分が利用さ

図3 修正KHモデル



(出所) 筆者作成

表2 「修正KHモデル」に基づく3つのマトリックス

A) 「a.利用者」×「b.利用対象」

a. 利用者	他 部 門					
	幹 部 職 員					
	管 理 職					
	施策担当者					
		評価活動	基本情報	実施状況	実施結果	判断・提言
		b.利用対象				

B) 「a.利用者」×「c.用途」

a. 利用者	他 部 門				
	幹 部 職 員				
	管 理 職				
	施策担当者				
		施策理解	意思決定	説得・支持獲得	アカウンタビリティ
		c.用 途			

C) 「c.用途」×「b.利用対象」

c. 用 途	アカウンタビリティ					
	説得・支持獲得					
	意 思 決 定					
	施 策 理 解					
		評価活動	基本情報	実施状況	実施結果	判断・提言
		b.利用対象				

(出所) 筆者作成

れるのかを把握することができる。

この修正KHモデルは、政策評価における利用実態を分析するための理論的枠組みとして用いるものである。具体的には、表2に示した「利用者」「利用対象」「用途」の3次元と各次元の要素を念頭において調査をおこなうことにより、評価の利用実態を体系的に理解することにつながる。具体的には、表2の3つのマトリックス表のそれぞれの組み合わせがどのような状態であるかを確認することにより、行政機関における政策評価の利用実態の特徴を明らかにすることができる（しかも、定量的・定性的のいずれの方法にも適用が可能である）。付論では、実際にこれらのマトリックスを用いて、政策評価の実態を試論的に検討してみ

る。

また、修正KHモデルは、政策評価を実施する行政機関自身にとっても有用である。このモデルの体系を念頭に置くことにより、利用面に注目して評価制度を設計することが容易になる。あるいは、このモデルを用いて自己の組織における政策評価の利用状況の実態や特徴を把握することにより、利用面の課題や問題点を発見することにつながる。

ここまでで本稿の目的をほぼ達成したことになるが、以下では付論として、政策評価制度における評価の利用状況を試論的に検討した結果を示したい。そのねらいは、修正KHモデルの利用方法を具体的に示すことと、政策評価の利用状況につ

いて仮説構築の基礎とすることの2点である。

付論．政策評価の利用状況の分析

(1) 分析の方法

前項の最後で述べたように、修正KHモデルの利用方法を示すことと、政策評価の利用状況についての仮説構築の基礎とすることの2点をねらいとして、修正KHモデルを用いて、国の政策評価における利用状況を分析する。直接的に利用するのは、直前に示した3つのマトリックスである。

それぞれのマトリックスで、各マス目の組み合わせについて評価の利用度合いを検討する。本稿では、「評価の利用」を「評価を実施する際にあらかじめ想定していた目的に用いること」と定義している。そこで、マトリックスの各マス目について、評価の利用が想定される組み合わせであるかどうかをまず検討することにした。

評価の利用が想定される組み合わせでは、(本稿の定義による)「評価の利用」が発生しうるので、各マス目の組み合わせについて実際の評価の利用度合いを検討する。利用度が「高い」組み合わせには「◎」、利用度が「ある程度高い」組み合わせには「○」を表示する。評価の利用が想定されるものの、実際の利用度が低い場合はマス目を空欄とする。一方、評価の利用が想定されない組み合わせ(つまり、本稿における「評価の利用」に該当しない組み合わせ)のマス目は黒く塗りつぶして示すことにした。以上から、マトリックスのマス目は、「◎」「○」「空欄」「塗りつぶし」のいずれかで表示されることになる。

各マス目を「◎」「○」「空欄」「塗りつぶし」のどの表示にするかは、筆者が長らく観察してきた文部科学省における政策評価制度の実施状況³を中心にして、筆者の知見やこれまでの観察結果に基づき、定性的かつ直観的に判断をおこなった。先述したように、今回はあくまで試論として分析をおこなうものであり、体系的な調査に基づく厳密な分析結果ではないことをあらかじめ断っておく。

以下では、それぞれのマトリックスについて分析結果を示す。

(2) 「a. 利用者」×「b. 利用対象」

表3は、「利用者」と「利用対象」のマトリックスである。前述したように、各マス目は評価の利用度に応じて「◎」「○」「空欄」「塗りつぶし」のいずれかの表示となっている。また、評価の利用が期待される組み合わせのマス目は太枠で囲んだ(太枠で囲んだ領域を「評価利用期待領域」と呼ぶことにする)。

表中で「施策担当者×評価活動」と「施策担当者×基本情報」の組み合わせには「◎」を表示した。これは、評価を実施する施策担当者自身が、評価活動そのものや施策の基本情報を利用しているという実態を示している。

施策担当者は、評価活動や施策の基本情報を「利用している」という意識は希薄かもしれない。だが、評価の一連の過程に関与することや、施策に関する基本的な情報から得られる知見は、施策担当者にとって利用価値があり、実際にこれらを利用することにより効用が発生していると判断した。後述する表4の分析結果を先取りすれば、施策担当者のこうした行為は、自身の担当する施策に関する理解(「施策理解」)を深めることにつながっているというのが本稿における解釈である。

一方、「施策担当者×実施状況」と「施策担当者×実施結果」の組み合わせは「○」とした。これらの組み合わせは、本来、評価の利用が最も期待される組み合わせである。だが、各府省が施策の実施状況や成果を把握するために設定している評価指標の数・質やその分析が不十分であるという現状認識に基づき、敢えて利用度を低く判断した⁴。

次に管理職は、実施状況、実施結果、判断・提言を利用している。利用しているといっても、施策担当者が作成した評価結果や判断を確認し、追認するという関与が中心だとみられる。評価結果等を積極的に利用しているわけではないことから「○」を表示するに留めた。

幹部職員や他部門は、評価のどの対象もほとんど利用していないと判断したことから、マス目はすべて空欄とした。ただし、この見立てには異論があるかもしれない。評価の実施結果や判断・提言を「見ている」幹部職員は少なくないであろうし、他部門でも例えば財政担当課では、予算要求

表3 「a. 利用者」×「b. 利用対象」のマトリックス

		評価の利用が想定されない組み合わせ		評価の利用は想定されるが、評価が 利用されていない組み合わせ		評価利用可能領域（太枠内）	
a. 利用者	他 部 門						
	幹 部 職 員						
	管 理 職			○	○	○	○
	施策担当者	◎	◎	○	○		
		評価活動	基本情報	実施状況	実施結果	判断・提言	
		b.利用対象					
評価の利用度が「高い」組み合わせ		評価の利用度が「ある程度高い」組み合わせ		評価の利用が期待されるが、評価が 利用されていない組み合わせ			

（出所）筆者作成

書に評価結果が添付される場合が多いことから、評価結果を「参考にしてている」とは言えるかもしれない。こうした状況を「利用」の範疇に含めることも不可能ではないが、たとえ評価結果を見たり参考にしたりしていても、これらの主体の意思決定に影響を及ぼすような使い方ではないと判断したことから、表3では空欄で示すことにした（そもそもこれらの主体が評価を使っているのであれば、評価が使われていないことが大きな問題として指摘されることはなかったはずである）。

表中では、太枠で囲んだ「評価利用期待領域」のうち空欄となっているマス目が3か所あるが、これらの組み合わせにおいては、評価の利用が期待されるものの、実際にはほとんど評価が利用されていないことを示している。なお、幹部職員や他部門が評価活動そのものを利用することや、施策担当者が自身がおこなった判断・提言を利用することは想定しにくいことから、対応するマス目を「塗りつぶし」で表示した。

(3) 「a. 利用者」×「c. 用途」

次に表4は、「利用者」と「用途」のマトリックスである。各マス目の表示方法は表3と同様なので、説明は省略する。

先の表3では、施策担当者自身が評価の最も主要な利用者であり、評価対象の中でも評価活動と施策の基本情報の利用度が高いことが示唆されて

いた。表4では、施策担当者がどのような用途に評価を利用しているかを確認することができる。施策担当者について最も利用度の高い用途は「施策理解」である。つまり施策担当者は、自身の担当する施策についての理解を深めるために評価を「利用」しているということである。

国の行政機関が利用する評価シートの様式を見ると、施策や事業の基本的な情報（目的、内容、事業費等）を説明する欄が必ずといっていいほど設けられている。もちろんこうした記入欄は、他者（管理職、幹部職員、他部門、国民など）に対して施策や事業の概要を説明するために設けられているものであるが、実は、こうした記入欄に記述することにより、施策担当者自身が施策や事業の内容を改めて確認することができるという効用がある。また評価をおこなうために、施策に関するさまざまな情報を収集し、その結果を分析するという一連の評価活動は、施策担当者が担当施策についての理解を深め、これをより良く運営することにつながっていると考えられる。こうした判断に基づき、[施策担当者×施策理解] のマス目は「◎」で表示した。

施策担当者は、説得・支持獲得やアカウントビリティのためにも評価を利用している。ただし、意思決定の欄は空白とした。施策担当者が評価を利用して、自身の担当する施策に関して何らかの意思決定をおこなっている可能性も否定できない。とはいえ、施策担当者は評価を実施した結果

表4 「a. 利用者」×「c. 用途」のマトリックス

a. 利用者	他 部 門				
	幹 部 職 員				
	管 理 職		○	○	○
	施策担当者	◎		○	○
		施策理解	意思決定	説得・支持獲得	アカウンタビリティ
		c. 用 途			

(出所) 筆者作成

表5 「c. 用途」×「b. 利用対象」のマトリックス

c. 用 途	アカウンタビリティ				○	○
	説得・支持獲得				○	○
	意 思 決 定				○	○
	施 策 理 解	◎	◎	○	○	
		評価活動	基本情報	実施状況	実施結果	判断・提言
		b. 利用対象				

(出所) 筆者作成

に基づいて意思決定をおこなっているというよりも、あらかじめ意思決定していたことを説明するために評価を実施している面が強いと判断した。

管理職は、意思決定、説得・支持獲得、アカウンタビリティのために評価を利用していると考えられる。ただし、いずれについても積極性が高い利用の仕方とはいえない。

表3では、幹部職員と他部門は評価を利用していないことが示唆されていたが、その結果を踏まえて、表4でも（塗りつぶされているマス目は除き）両者のマス目は全て空欄で表示した。

ところで、アカウンタビリティのための利用については、補足が必要である。表4によれば、施策担当者と管理職が評価をアカウンタビリティのために利用している。ただし、これは施策担当者や管理職が評価結果を携えて特定の人物や場所に出向き、評価結果について説明したり、その結果について責任を取ったりしていることを必ずしも意味していない。

むしろ実態としては、評価を実施することがルーティンに組み込まれており、しかるべき時期に施策担当者や管理職が評価シートを作成すれば、

その結果がとりまとめられて組織内部で共有されたり、対外的に公表されたりすることで、アカウンタビリティを果たしたものとみなされる。つまり、一定のアカウンタビリティを果たすことがあらかじめ制度的に担保されている。したがって、アカウンタビリティ目的に利用するために、施策担当者や管理職に特別な意識や努力が必要なわけではない点には留意が必要である。

(4) 「c. 用途」×「b. 利用対象」

最後に表5は、「利用者」を捨象して「用途」と「利用対象」の関係を示したマトリックスである。表5は、利用対象と用途の関係を独立して示しているわけではなく、その背後には表3と表4の分析結果が介在している。

表5では、太枠で囲まれた「評価利用期待領域」のすべてのマス目が「◎」または「○」で埋められている。利用度の高い組み合わせは「施策理解×評価活動」と「施策理解×基本情報」である（いずれも「◎」）。表3、表4の結果と併せて考えると、評価活動や施策の基本情報を主に利用しているのは施策担当者であったことから、施策担当

者が施策理解のために評価活動や基本情報を利用していることが改めて確認できる。

一方、意思決定のために評価の実施結果や判断・提言は利用されているものの、その利用度は高いとは言えない。表4において、意思決定のために評価を利用しているのは専ら管理職であったから、管理職がこれらの材料を積極的に利用していないことを反映した結果である。

5. まとめ

付論で示した内容は、厳密な調査や分析に基づくものではないことから、あくまで現状についての試論的な検討結果に過ぎない。ただし、国の政策評価における利用状況の傾向を示唆するものとして、一考する価値は認められるであろう。

これまでの検討結果を整理すると、次のとおりである。

第1に、政策評価においては、施策担当者が最も主要な利用者であり、施策担当者が自身の担当施策について理解を深めるという用途が中心になっていることが示唆された。政策評価制度では「評価結果を政策に反映」させることを重視しているにもかかわらず、実態としては、評価の「結果の利用（results use）」よりも「過程の利用（process use）」が優勢となっている可能性を示唆している。

第2に、管理職は意思決定、説得・支持獲得、アカウンタビリティの用途で評価を利用しており、管理職の役割に照らして、それは妥当である。ただし、それらの用途に利用しているとしても、管理職が積極的に評価を利用しているとは言えない。現実的には、施策担当者がおこなった評価の結果を確認したり、追認したりすることが、最も一般的な利用形態だと考えられる。

第3に、幹部職員は評価をほとんど利用していないと考えられる。その一因として、現状の政策評価の内容が、幹部職員にとって有用性が高いものではないことが挙げられる。

第4に、他部門による評価の利用もほとんど見られない。例えば、財政担当部門が予算要求の査定のために評価結果を利用するという可能性が想

定されるが、それがほとんど実施されていないということである。その理由はさまざまであろうが、評価情報に対する信頼性が低いことが、他部門による評価の利用を妨げている可能性がある。

第5に、アカウンタビリティを果たすための評価の利用は、制度的に担保されており、評価者は特別な意識や努力を必要としない。言い換えれば、組織としては全体的なパフォーマンスについて一定のアカウンタビリティを果たしていることになるが、個別の施策や部門のレベルでは、十分にアカウンタビリティを果たしているとは言えない可能性がある。

以上の検討結果を総合すると、「評価の利用」に対する一般的なイメージに反して、国の政策評価においては、極めて限られた利用者（施策担当者と管理職）が、限定された用途（施策理解が主で、意思決定等への利用は消極的）に評価を利用しているという可能性が示唆された。これは筆者の知見や観察に基づくものであるため、現実の投影としては幾分誇張を含んでいるかもしれない。とは言え、意思決定への利用が不十分であることや幹部職員や他部門による評価の利用がほとんど見られないことは、今後の評価制度のあり方を考える上で重要な問題提起としてとらえることができるであろう。

最後に、本稿では評価の利用を理論的に解明したことにより、評価の利用を客観的に把握・分析するための枠組みを構築することができた。今回示した枠組みをさらに洗練させるとともに、今回は試論レベルに留まった政策評価の利用の実態に深く踏み込むことを今後の課題としたい。

注記

- 1 例えば、筆者が2008年に静岡県内の一部の団体（県、市、町）の地方議員（対象者261名）を対象として実施した調査では、「評価結果等を頻繁または定期的に見ている」と回答した議員は、全体の10%弱であった。
- 2 「施策」と表記するが、施策とその下位の階層の「事業」を含む。以下に登場する「施策」も同様。
- 3 筆者は同省の政策評価に関する有識者会議委員を長年務めているため、同省の政策評価制度の内容と

その実施状況を継続的に把握してきた。

- 4 評価指標の数・質や分析が不十分であるということは、施策の実施状況や実施結果に関する評価結果が十分な質を備えていないことを示唆している。評価を実施した本人である施策担当者は、その事実気づいているはずであるから、それらの評価結果を積極的に利用しようとはしないであろう。

参考文献

- 佐々木亮（2010）『評価理論：評価学の基礎』、多賀出版
- De Lancer Julnes, P. & Holzer, M. (2001). Promoting the Utilization of Performance Measures in Public Organizations: An Empirical Study of Factors Affecting Adoption and Implementation. *Public Administration Review*, 61 (6), 693-708.
- Fitzpatrick, J. L. et al. (2004). *Program Evaluation: Alternative approaches and Practical Guidelines* (3rd Ed.). Allyn and Bacon, Boston:MA.
- Kirkhart, K. E. (2000). Reconceptualizing Evaluation Use: An Integrated Theory of Influence. *New Directions for Evaluation*, 88, 21-33.
- Patton, M. Q. (1996). *Utilization-Focused Evaluation: The New Century Text* (3rd Ed.). Sage Publications, Thousand Oaks: CA.
- Perrin, B. (1998). Effective Use and Misuse of Performance Measurement. *American Journal of Evaluation*, 19 (3), 367-379
- Weiss, C. H. (1988). *Evaluation* (2nd Ed.). Prentice Hall, Upper Saddle River: NJ.
- Weiss, C. H. (1998). Have We Learned Anything New About the Use of Evaluation? *American Journal of Evaluation*, 19 (1), 21-33.

(2013.9.19受理)

Evaluation Use in The Policy Evaluation System: Cocentralizing “Evaluation Use” and a Tentative Analysis of Present Practice

Hiraki Tanaka

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

Abstract

More than eleven years have passed since the central government organs in Japan started to implement policy evaluation. In recent years, many commentators claim that the evaluation conducted under the policy evaluation system has been under-utilized. Because the definition and the concept of “evaluation use” have remained unresolved, the present state of the “actual” evaluation use has not been well understood. Based on such recognition, this article tries to build a framework for empirically analyzing evaluation use of the central government organs through clearly conceptualizing the idea of “evaluation use.” The Kirkhart’s theoretical framework, which this article consulted for theoretical reference, is a model based on a societal viewpoint to understand the serial process in which the effect of program evaluation is diffused. In this article, by modifying the Kirkhart’s model, a new theoretical framework with an organizational focus is proposed of which purpose is to analyse the evaluation use under the policy evaluation system in Japan. A tentative analysis of the current state of evaluation use by the central government organs revealed that the users of evaluation are mainly the public servants in charge of individual programs and their superiors, and that the purpose of use is predominantly for understanding their own programs in charge. The theoretical framework proposed in this article is an effective tool for empirically analyzing evaluation use in public organizations. Further analyses using the framework are anticipated.

Keywords

Policy Evaluation System, Government Policy Evaluations Act (GPEA) , evaluation use,
utilization-focused evaluation

【研究論文】

政策評価とアカウンタビリティ ～法施行後10年の経験から～

南島 和久

神戸学院大学

najima@law.kobegakuin.ac.jp

要 約

行政機関が行う政策の評価に関する法律（平成13年法律第86号。以下「政策評価法」という。）が施行されてから10年が経過する。本稿では、同法において掲げられた政策評価制度の目的のひとつであるアカウンタビリティについて、「この10年の経験で行政機関のアカウンタビリティは向上したのか」という点を問いたい。

本稿の結論は、政策評価制度の導入により行政機関の透明性は向上したものの、結果的にアカウンタビリティが向上したとまではいえないというものである。ただし、ここで重要な論点は、「アカウンタビリティが向上しなかった」という結果ではなく、「なぜアカウンタビリティの向上がかなわなかったのか」という点の分析や「そもそも日本の行政機関にとってアカウンタビリティとは何か」という点にこそある。

本稿は、日本の行政機関における政策評価とアカウンタビリティとの関係についての、行政学の立場からの理解を示すものである¹。

キーワード

アカウンタビリティ、権力分立、政策評価法、政策と管理の論理

1. アカウンタビリティと学問分野

政策評価に貢献する学問分野は多種多様である。それは「応用社会科学」(applied social science)ともよばれる政策評価の特徴に由来している。

政策評価に関係する代表的な学問分野は、経済学、経営学、政治学、社会学などである。また、その手法論については、会計学、OR (Operations Research)、ゲーム理論、システムズ・アナリシス、統計学、社会調査、文化人類学などでの蓄積

が重要とされている。もちろん、その他個別政策分野の貢献も重要である。教育、福祉、ODA（政府開発援助）、外交、防衛、研究開発、土木、都市計画などの分野別の知識は、政策内容の面での充実に寄与するものである。

こうした多彩な諸学問分野のなかで、政策評価に「アカウンタビリティ」を持ち込もうとしてきたのは、とくに「会計学」と「行政学」であった。その特徴を要約すれば、いずれも「大規模組織のコントロール」、あるいはトップマネジメントを支える「ミドルマネジメント」に関心を寄せてい

ると表現できる。なお、「ミドルマネジメント」について補足しておく、経営学では「経営管理」、行政学では「行政管理」と表現されるが、いずれも英語では、“administrative management”である。

最初にアカウンタビリティという言葉の問題からはじめたい。先に「会計学的なアカウンタビリティ」論を概観し、その上で「行政学的なアカウンタビリティ」論を確認しよう。

(1) 会計学的なアカウンタビリティ

まず「会計学的なアカウンタビリティ」についてである。

会計学は「アカウンティング」、すなわち「企業会計」を主な対象とし、大きくは財務会計論と管理会計論（management accounting）とに分かれる。政府の政策評価活動と関係し影響を与えているのは、これらのうち「管理会計論」である。管理会計論は意思決定を支える経営情報の集約に力点があり、ここには業績測定（業績評価）やバランス・スコア・カードといった手法論などが含まれている。

アカウンタビリティ論は、財務会計論と管理会計論をあわせた会計学全体のテーマであり、会計学では「会計責任」という訳語が当てられている。その中心にあるのは貸借対照表や損益計算書などの財務諸表である。

会計学的なアカウンタビリティとは、究極的には企業組織の所有者たる株主に向けられる、株主の投資を支えるための信頼確保の道具立てとしての責任構造の議論である。したがってこの概念は、狭義には財務諸表を中心とした財務系のマネジメントを意味している。そしてそこには、外部監査の精度をあげるための内部統制の構築、とくにコンプライアンス構築による企業価値の担保、あるいはより積極的に企業の社会的責任の充実などが含まれる。

日本の行政機関におけるアカウンタビリティの概念、すなわち行政機関自身による「国民への説明責任」はこれと何が異なるのか。会計学的なアカウンタビリティは、一方で経営情報の「説明」を中心とする。その点に注目するならば行政機関で用いられている「説明責任」はこの翻訳といえよう。しかしこの両者において決定的に異なるの

は、「問責の外在性」や「客観性」の構造である。控えめにいったとしても日本の行政機関におけるアカウンタビリティの概念においては、「外部から行政組織をコントロールする」イメージはさほどつよくない。

このアカウンタビリティの論理構造や責任の体系をめぐる民間企業との異同の論点についてはまたのちほど触れたい。その前に、会計学的なアカウンタビリティをふまえつつ展開した行政学的なアカウンタビリティの概念を経由しておこう。

(2) 行政学的なアカウンタビリティ

現代行政学はアメリカ由来の学問であり、その初期には経営学の影響を色濃く受けてきた。とりわけ、政治と行政とを峻別し、行政を究極的に規律する価値として「能率」(efficiency)をおく点に、行政学の学問的特徴はあると考えられてきた。

行政学的なアカウンタビリティ論の内容に入る前に、アメリカ行政学の全体像をおさらいしておきたい。西尾勝（2001：28-35）は「統制の規範」「分離の規範」「協働の規範」という3つの表現でアメリカ行政学の理論的特徴を整理している。この議論はアメリカ行政学の史的展開を大掴みにする枠組みである。

まず「統制の規範」は、行政が政治のコントロールの服すべきだという規範的な主張を含むものである。これは、行政の民主的コントロールにかかわる議論であって、アンドリュー・ジャクソン第7代大統領の時代に展開した「ジャクソニアン・デモクラシー」に典型的に見られるような、猟官制（spoils system）の伝統に基づく議論である。もっともジャクソン流の猟官制は、次第に政治と行政の融合・癒着、ついで腐敗・汚職の温床とも目されるようになっていった。

この「統制の規範」の文脈からアカウンタビリティを論じるなら、「政官関係」がその中心に位置するだろう。理念的な政治責任の姿を結果責任におくなら、行政官の責任はこれに対する制度的なく説明・弁明と表現できる。そこには、政治に傳く行政の姿、政治からのコントロールに服する行政の在り様がある。

つぎの「分離の規範」は、政治と行政を分離し、行政の独自の領域、すなわち行政の政治的中立性

を獲得しようとするものである。政治と行政の分離（学説史上、「政治・行政分離論」という）については、資格任用制を掲げた連邦公務員法たるペンドルトン法（1883年）や、ウッドロー・ウィルソンの「行政の研究」（1887年）などがその嚆矢とされている。政治・行政分離論の背景には、資本主義の拡大とそれに伴う政治社会の歪み、とくに腐敗などの猟官制のマイナスの影響、社会の近代化に伴い行政にも技術的合理性や能率が求められるようになったことなどがあげられている。

分離の規範の積極的意義は、政治の領域から行政の領域を分離し、そこに企業経営の理念とされた「能率」規準をあてがい、行政組織の政治的中立性を図ろうとした点に求められる。ここでのアカウンタビリティ概念は、会計学的なアカウンタビリティの議論のうち、とくに節約と能率の摂取が目进行く。行政組織の拡大とともに大規模組織をコントロールするための「行政管理」の概念が導入されたのもこの時期のことだが、とくに財務会計系のアカウンタビリティ概念の政府部門における発達は、この分離の規範の作用に負うところが大きい。

最後の「協働の規範」は、政治と行政のそれぞれの役割分担のもと、政治・行政のそれぞれの責任を確立し、そのうえで政治と行政との間に協働関係を構築しようとするものである。ニューディール期以降の福祉国家（welfare state）の時代には、能率論への偏向や党派的性格への批判が先鋭化した。とくに政策面での批判は大きく、政治と行政の連続面、役割分担、新たな責任体制などが強調されるようになった。なお、分離の規範から協働の規範へのアクセントの移動は学説史上、「政治・行政分離論から政治・行政融合論へ」とよばれている。

政治・行政融合論では、政治・行政分離論での能率規準にかえて、専門集団としての行政組織の機能面と政治との役割分担に注目した議論、すなわち「行政責任論」が注目を集めた。行政責任論の焦点のひとつは、専門職能の職務に関する結果・成果論として行政責任を理解する点にあった。ここにおいて、あらためて政治と行政との間の役割と責任の再整理が模索され、のちに触れるようにそのなかでアカウンタビリティ概念も論じられていった。

公的部門においてアカウンタビリティ概念に顕著に注目が集まったのは、この「協働の規範」以降の時期である。それは行政活動の範囲と規模、行政機関の役割が急速に拡大と、それに伴う政治による行政のコントロールが作動しなくなる行政国家の時代における、「代議制の危機」を象徴する議論でもあった。

2. 政府のアカウンタビリティ

「協働の規範」の時代以降、アカウンタビリティの概念は、「代議制の危機」を前に注目を集めるようになった。現代政府におけるアカウンタビリティ概念の展開については、大きく3つの議論をふまえたい。第1に1940年代の行政責任論争である。第2に1960年代から1980年代にかけてのプログラム・アカウンタビリティの展開である。第3に1990年代以降の財政構造改革論＝NPMとしてのアカウンタビリティの議論である。

(1) 行政責任論争

第1のアカウンタビリティに関するトピックは行政責任論争である。その内容は日本の標準的な行政学の教科書にはかならず紹介されている。

行政責任論争とは、1940年代に展開したカール・J・フリードリッヒとヘルマン・ファイナーとの間に展開した論争を中心とする議論のことをいう。フリードリッヒは「公共政策と行政責任の特質」（Friedrich 1940）という論文において、代議制の危機の時代における行政官の責任を論ずるにあたり、その専門能力に注目しつつ、行政官の民衆感情への応答性（政治的責任）を強調した。それが口火を切ることとなった議論の端緒である。

ファイナーはすぐさまこれに反論し、同年の「現代政府における行政責任」（Finer 1940）という論文において、フリードリッヒの立論を行政官の責任感＝道義的義務感（moral obligation）論であるとし、そこに欠落している論点、すなわち、議会に対し行政機関がアカウンタブルであるべきだという点を強調した。この論文においてファイナーはとくに行政の政策に関する議会への説明責任を強調した。

この際、ファイナーは自説を「X is accountable for Y to Z」という公式によって表現した。行政(X)が議会(Z)に対し、その活動内容(Z)についてアカウントブルであること(説明・弁明しうること)を求めたこの議論は、フリードリッヒの立論のままではナチス・ドイツのような権力の暴走(全権委任法に象徴される)を止められないようなものではないかという指摘とともに展開したものである。それは、特定の権力の暴走を防ぐためのチェック・アンド・バランスを主眼におく古典的な権力分立論の再生でもあった。ファイナーの反論において強調されたアカウントビリティはその後、行政学においては議会への<説明・弁明>を求めるフォーマルな「制度責任」「法的責任」として紹介されていった²。行政国家の時代において、それはさらに、政治以外の政府部内でのコントロール方法の模索へと発展していった。

これらの論争に加え、政治以外の政府部内での行政のコントロール方法についてはチャールズ・ギルバートが示した枠組み(Gilbert 1959:382)が著名である。<ギルバートの構図>とよばれるこの議論では、「外在」「内在」と「公式」「非公式」で分かたれる4つの象限に行政へのコントロールが整理された。その議論の内容は、「外在・公式」の象限には議会や司法によるコントロールが、「外在・非公式」には評論家や市民などによるコントロールが、「内在・公式」には階統型の組織内部のコントロールが、「内在・非公式」にはモラルなどの行政内部の規範がそれぞれ整理されるというものであった。

ギルバートの議論はファイナーの流のアカウントビリティ論を拡充してみせた点に妙味があったといえるだろう。すなわち、ファイナーが論じた「外在・公式」の行政責任のほかに、現代政府のアカウントビリティは、それが他の三象限にも拡散する可能性を示唆したのである³。それはまさに代議制の危機と表裏の関係におかれた論点であった。

ところで、これらの責任領域のうち「内在・非公式」の領域については、外部から責任を調達することができないものであった。アカウントビリティ論は、会計学説でみたように、「問責の外在性」とそれによって調達される「客観性」を基軸

とするものである。他方で「内在・非公式」の責任領域(「組織内部」の「非公式」な責任のあり方)では、行政官のモラルへの依存が過大となり、ファイナーが指摘していたように、「外在性」や「客観性」の論点が希薄化してしまいかねない。

それでは、「『内在・非公式』の責任領域が軽視されてもよいか」というとそうではない。なぜなら、行政の高度化・専門化はまさにこの「内在・非公式」の領域において展開するのであって、「内在・非公式」のコントロールをどのように担保するのかという問題がまさに行政責任論の本質だからである。フリードリッヒの論点提起もまさにこの点をとらえたものであった。

かりに、「内在・非公式」を除く他の三象限におけるコントロールが十分に展開しえない場合には、フリードリッヒが立論したように、この「内在・非公式」の領域のコントロールについての民主的コントロールへの期待が集まることになるだろう。それはひとことでいえば専門家集団としての倫理の問題(=プロフェッショナリズム)層の議論である⁴。現代政府の行政責任として、この「内在・非公式」の領域における行政責任論が重視されてきた理由はここにある。この意味においてフリードリッヒの論点提起はまさに正鵠を得るものであった。

この一連の行政責任論争では、そこにアカウントビリティとレスポンシビリティという2つの責任論があること、およびこれらの一部がコインの表裏の関係におかれていることを示した点に意義がある。そして、公的部門におけるアカウントビリティの概念や政策評価論はこの2つの行政責任にまたがる課題を抱えているのである。

(2) プログラム・アカウントビリティ

第2のアカウントビリティに関するトピックは、1960年代以降のプログラム・アカウントビリティの概念についてである。プログラム・アカウントビリティの展開の前提には、「プログラム」という概念への注目がある。プログラム・アカウントビリティへの注目に関し、2つの文脈を指摘しておきたい。

第1の文脈は、1921年の予算会計法に基づいて設置された議会付属機関であるGAO(米国会計検査院)のアカウントティングへの接近と摂取であ

る⁵。GAOは連邦予算の監査に際し、戦後、会計士などの専門職の採用を熱心に行ってきたという経緯があった（平井 2012）。このことにより、GAOを媒介として、「政府のアカウンタビリティ」（Government Accountability）が強調されるという結果がもたらされた。会計学的なアカウンタビリティ（＝「会計責任」）が会計専門職を通じて政府に摂取されたのである。なお、GAOは連邦政府の機関ではなく、連邦議会付属機関であって権力分立上も外部性を主張しうる位置にあったこと、GAOが確保しようとしたものが外在的な客観性であったこともここに影響していた。

会計学的なアカウンタビリティを摂取したGAOの評価活動は、1970年代にはPPBSの合流により、「監査から評価へ」とよばれる変貌を経験した。この場合の「監査」は会計監査のことであり、「評価」はプログラム評価のことである。このような力点の移動は、後年、GAOの“Accounting”から“Accountability”への名称変更にも反映されたものである。GAOは設置から2004年までは“General Accounting Office”を名乗っていたが、2004年以降は“Government Accountability Office”へとその名称を変更したことは周知のとおりである。

第2の文脈は、「プログラム予算」（program budget）に始まる「プログラム」を基礎とする政府政策の考え方である。プログラム予算とは、プログラムと予算の単位とを一致させるという考え方のことをいう。それは政府の予算管理上も重要なことだと考えられてきた。

このプログラム予算は1963年連邦予算に一部導入された。当初、プログラム予算はランド研究所の経済学者、OR研究者などの意見を参考にして軍事予算に試みられた。その基本的な考え方はその直後のPPBS（planning, programming and budgeting system）へと継承された。PPBSではプログラム予算はさらに計画（planning）とも連携することとされていった。

いずれにせよ重要であったのは行政活動の括りである「プログラム」という単位への注目であった。その中心には、プログラムの「有効性」（effectiveness）（＝成果・効果）への注目が絡んでいた。「プログラム」への注目は、PPBSを介しつつ、議会主導のGAOのプログラム評価へと流

れ込んだ。議会の側としても、ウォーターゲート事件などに絡み、その信頼性の危機に瀕していた。プログラムとその有効性への傾斜は、GAOが媒介者となり、会計責任に端を発するアカウンタビリティの概念に接木され、議会を媒介とした責任確保論へと展開していったのである（山谷 1988b; 益田 2010）。

GAOの行うプログラム評価の意義は、議会の意向をふまえつつ、行政活動＝政策実施をふまえた調査・分析等を行う「事後評価」（ex-post evaluation）を行うことにあった。プログラム評価の特徴は、政策の実施後に学術的な手法を用いた調査を行い、それを専門職員や外部のシンクタンクの活動が担い、信頼性と客観性に基づく調査結果を産出する点にあった。こうしたプログラム評価は、のちに政策評価のひとつのモデルとなった。それが「政府のアカウンタビリティ」とよばれたものの一断面であった。

（3）業績測定

第3のアカウンタビリティに関するトピックは、1990年代以降の財政構造改革論＝NPM（New Public Management）の議論である。1990年代のクリントン民主党政権の時代には、行財政改革がひとつの政治課題となった。このため、連邦政府の予算を省庁自らが管理するGPRA（Government Performance and results act）が導入された。GPRAはオバマ民主党政権において大幅な改定を受けつつ現在も継続されているところである。

GPRAにおける評価手法は「業績測定」（performance measurement）とよばれるものであった。アーバンインスティテュート（UI）が中心となって開発したこの手法は、民間企業において普及している目標管理・成果主義を公的部門にも導入可能としたものであった（Hatry 1999）。その基本的な考え方は、オズボーンとゲーブラーの『行政革命』（Osborne and Gaebler 1993）にもよく示されているように、企業モデルの手法の強調にあった。

かつて全連邦政府機関に導入しようとして失敗したPPBSの反省をふまえ、業績測定は大幅に簡素化され、行政機関内部で安定的に運用しうるように工夫されたものであった。PPBSではORやシ

ステムズ・アナリシスといった高度な分析手法の導入が企図されたが、それは行政実務にとっては過剰な負担でしかなかった。こうしたPPBSやプログラム評価の際に強調された準学術的側面を捨象したところに業績測定は成立をみたのである。

「アントレプレナー型（企業家的な）政府」を標榜するクリントン政権の方針も相俟って、業績測定は政策評価のひとつのモデルの地位を獲得するに至った。業績測定は行政機関にとって導入の敷居が低く、また運動論的でもあった。そこで繰り返し喧伝されたのは、行政官の「意識改革」であった。先のチャールズ・ギルバートの行政責任の構図でいえば、「内在・非制度」の行政責任領域を中心に、民間企業経営者のような経営感覚を具備する行政官像が求められたのである。

「内在・非制度」領域における経営感覚としての行政責任の強調は、アカウンタビリティの議論にひとつの変化をもたらした。それは権力分立論としてのアカウンタビリティ論、すなわち「外部」ないし「公式」に調達されるアカウンタビリティ論から、「内在・非制度」領域へのアカウンタビリティ概念のシフトであった。この領域の行政責任を強調したものが、NPM論者であった。

業績測定の理論的背景にはNPMがあるのだと指摘される。NPMは「企画立案」と「執行」の分離、分離した執行の領域の「効率化」、それを実現する手段としての「業績測定」、そしてとくに目標の「定量化」を強調するものである。また、行政機関に企業モデルのマネジメントを求め、組織経営の健全性を要求するものである。それはかつての政治・行政分離論やフリードリッヒ流の行政責任論を彷彿とさせる。業績測定において、「マネジメント」「マネジメント・サイクル」が強調され、ときに「行政『経営』」という表現が登場するのはこのような文脈で理解できるだろう。

ところで、業績測定との関係で強調されたアカウンタビリティは、政府サービスの水準を客観的な数値で表現すること（定量評価）であった。日本でもこうした業績測定、とくに指標の定量的表現が普及した。そして、それこそがアカウンタビリティであると観念され、その邦訳としての「説明責任」が定着してきた。

ここで論点を示しておきたい。政府で使われる

日本語の「説明責任」としてのアカウンタビリティ概念においては、英語で表現されるところのアカウンタビリティ概念に含まれる「会計責任」「法的責任」「制度責任」との関連性が希薄となってしまうている。日本語と英語のボキャブラリーの違いといってしまうえばそれまでなのかもしれないが、その一部だけが強調されることで概念の伝統やイメージの広がりが捨象されてしまっているように思われる点をここでは指摘しておきたい。

3. アカウンタビリティは向上したか？

日本の政策評価やアカウンタビリティは、最後に述べた業績測定レジームの圧倒的影響のもとで展開したものであった。それは日本の評価ブームが1990年代の後半から2000年代にかけてのことであったこと、行政機関の内部評価としての業績測定が普及したことなどを背景としていた。その裏側で、公共部門におけるアカウンタビリティの伝統に寄与してきた「権力分立の確立」「政策実施過程をふまえた議論」「社会科学の諸手法を用いた政策の調査・分析」「プログラムの観念の重視」「外部評価による中立性・客観性の確保」などは十分に重視されてはこなかった。このような視点から10年の経験を振り返ってみよう。なお、すでに数多くの著作で評価史については描写されているので、ここではアカウンタビリティに絡む論点のみを追いかけたい。

(1) 制度設計の段階

国の行政機関に政策評価が導入されることとなったのは、1997年に出された行政改革会議の「最終報告」以降である。ただし、行政改革会議では政策評価については十分に議論されておらず、その実質的な内容は法整備の段階や総務庁に設置された「政策評価の手法等に関する研究会」において詰められることとなった。このような背景をもつ日本の政策評価制度の導入時の論点として以下の5点を指摘しておこう。

第1に政策の企画立案と実施機能が分離され、実施機能の評価の議論が先行したことである。中央省庁等改革においては、内閣府の設置と省庁の

大括り化、独立行政法人制度の創設などが目玉であった。評価論は、まずこのうちの独立行政法人制度との関係で議論された。

独立行政法人評価では業績測定型の評価が採用されることになっていた。また、独立行政法人制度は典型的なNPMの制度といわれるように、企画立案と政策実施を分離し、実施部門の「効率性」をその中心におくものであった。さらに、独立行政法人の予算＝運営費交付金は業績・成果と連動するものとされ、評価は事実上の査定の仕組みとして整理された。その結果、独立行政法人評価では「有効性」よりも「効率性」の方が重視されることとなった。それは実施機関の評価として考えれば当然の帰結であった。そしてこうした評価のあり方が公的部門における評価のモデルとなっていった。

第2に実施機能の評価とともに国の行政機関への評価の導入に際しては「企画立案の評価」が注目されたということである。しかし、先にも触れたように、その具体的方法に関する行政改革会議での議論は十分なものではなかった。その後、「政策評価の手法等に関する研究会」においても「企画立案の評価」に関する議論は収斂することはなかった。そして、最終的に「事業評価方式」「実績評価方式」「総合評価方式」の3つの評価方式が併記されることとなり、その選択も含めて各行政機関のマネジメントに委ねられることとされた。もっとも、「公共事業」「研究開発」「政府開発援助（ODA）」といった多額の資金を要する事業等については例外的に厳しく管理されることとなった（義務付け三分野）。この義務付け領域にはのちに「規制」と「租特」（租税特別措置）が加わることとなる。

「企画立案の評価」がひとつの評価手法に収斂せず併記型となった点に関しては、会計学、行政学、経済学がそれぞれが念頭におく政策評価論が一致していなかったという背景があった。当初、評価を介して「共通言語」を創出することが強調されていたが、すくなくとも学問の壁を越えてこれが実現することはなかった。

また、外部評価よりも行政機関内の内部評価が重視され、府省側のマネジメントのもとに整理された点も特筆しておきたい。現行の政策評価制度

導入の直前期には民主党からは「日本版GAO構想」も打ち出されていた。だが、こうした外部評価については三権分立に反するとして否定されることもあった。あるいは、行政機関外部よりも内部の方が情報の量も質も充実しているのであるから、まずは内部評価からはじめるべきといった議論も行われていた。

第3に府省の自律性に関する論点である。中央省庁等改革においては、内閣府に対し各省より一段高い位置から「総合調整」を行うこととされ、各府省側では「政策調整」を行うことが期待された。これらはそれぞれ「垂直的調整」「水平的調整」ともいわれた（今村 2006; 岩田 2010）。内閣府設置法、国家行政組織法に書き込まれた「自ら評価」は、客観的な政策効果の検証をふまえて、これらの調整を行い、「最終報告」がいうように、セクショナリズムの是正へと向かうことが期待されていたものであった。

ここで強調しておきたいことは、内閣府が担うこととされた「総合調整」、各省間の「政策調整」、そして政策効果をとらえるための「政策評価」がもともと同じコンセプト（＝セクショナリズムの打破と総合調整）を有していたという点である。しかしそれらが相互に有機的連携をもつことはほとんどなかった。要するに制度設計の通りにはいかなかったということである。

第4に評価専担組織たる総務省が、従来の「行政監察」にかえて、「行政評価・監視」「統一性・総合性確保評価」を導入したことである。「行政評価・監視」は従来の行政監察の伝統を受け継ぐが、「統一性・総合性確保評価」は新たな外部評価の仕組みとして構想・設計されていた。いずれも府省横断的な、あるいは政府全体の統一的なテーマのもとで総務省が調査機関となって政策効果を把握せんとするものであり、それに基づく評価結果に関して総務省は勧告権を有するというものであった。

さらに総務省は評価専担組織としての評価として、「客観性担保評価」を行うこととされた。客観性担保評価は統一性・総合性確保評価とは異なり、各府省の政策評価の客観性をチェックするという一種の「メタ評価活動」であった。これらの新しく創設された評価制度は法律上の用語だが、

「政策の評価」とよばれた。

これら評価専担組織の評価活動は、設計当時からいくつかの問題点を抱えていた。そのうちもっとも深刻なものは、「政策効果の把握が十全に行いうるのか」という点であった。政策効果に関しては第一義的に個別の政策を所管する各府省が把握を行うこととされていたが、その責任が十分に果たされることはなかった。そこには政策効果の把握の手法の問題があった。社会科学の諸手法を動員した政策効果の把握という理念は、所管業務を主業務とするなかでの府省の自己評価活動では十分に展開できなかった。ここにはかつてのPPBSが悩んだのと同じ種類の論点が潜んでいた。

第5に法律として政策評価が整備されたということである。もともと政策評価制度は、内閣府設置法、国家行政組織法に埋め込まれる形で制度化されていた。各府省の政策評価活動は、制度的にはこの設置法レベルで実現しているものである。ところが、中央省庁等改革が行われ、各府省が政策評価活動に入った後に、政策評価制度が別途法制化されることとなった。それが「行政機関が行う政策の評価に関する法律」（政策評価法）であった。

政策評価法は各府省の設置法を前提としつつ、実質的にはその国会報告部分やその報告の仕様が義務化されているという建付けのものであった。その矛盾はそもそもの政策評価が各府省のマネジメントを前提に構想・設計されているという点にあった。すなわち、国会の関心を前提として設計されたものではなく、行政機関側のサブスタンスを书面化し、これを国会に報告するという構造になっていたのである。

(2) 制度運用の段階

国の府省に導入された政策評価は3つの目標を掲げていた。すなわち、「効率性」「成果重視」「説明責任」の3つである。政策評価法の第1条にこの文言は書き込まれた。その条文は以下の通りである。

(目的)

第一条 この法律は、行政機関が行う政策の評価に関する基本的事項等を定めることに

より、政策の評価の客観的かつ厳格な実施を推進しその結果の政策への適切な反映を図るとともに、政策の評価に関する情報を公表し、もって効果的かつ効率的な行政の推進に資するとともに、政府の有するその諸活動について国民に説明する責務が全うされるようにすることを目的とする。

政策評価の目的のひとつには本稿の主題でもある「国民に説明する責務」（＝アカウンタビリティ）が掲げられることとなった。そして、ここでの「国民に説明する責務」の内容は、「効率性」や「成果」（≡有効性）に関するものであった。政策評価法では「政策効果」（第3条）と表現されているものがこれである。まとめれば「効率性」や「政策効果」に関する事項が、行政機関から国民への「説明責任」として求められていたということができよう。

政策評価法の意義は、行政活動についての説明の合理性水準を引き上げようとする点にある。従来の行政活動においては行政機関からの説明は往々にして「法的根拠」や「権限」の議論に傾斜しがちであった。しかしながら、このような説明のあり方のままでは「費用対効果」などの経済合理性についての説明が不足していた。政府財政が逼迫している時代にあって、これは次第に共通認識となっていったように思われる。しかし、それは政策評価法が施行された後の制度運用の段階においても充実しえたといえるものではなかった。

ここまでの議論を念頭におきつつ政策評価制度とアカウンタビリティの10年を振り返ろう。論点として提起するのは以下の5点である。

第1に「評価書」についてである。政策評価法施行後においては、評価活動の具体的な内容は各行政機関が作成する評価書として表現されることとなった。評価書の内容は法第10条に示された以下の7項目であった。

(評価書の作成等)

第十条 行政機関の長は、政策評価を行ったときは、次に掲げる事項を記載した評価書を作成しなければならない。

一 政策評価の対象とした政策

- 二 政策評価を担当した部局又は機関及びこれを実施した時期
- 三 政策評価の観点
- 四 政策効果の把握の手法及びその結果
- 五 学識経験を有する者の知見の活用に関する事項
- 六 政策評価を行う過程において使用した資料その他の情報に関する事項
- 七 政策評価の結果

このうち重要なのは第三項から第六項までの4項目である。「すくなくとも当初は」という但し書きが必要かもしれないが、ここで前提・期待されていたのは、政策評価を行うにあたり、学術的な手法等を用いてその説明の合理性水準を向上させることにあった。すなわち、「法的根拠」「権限」の議論にとどまらず、所管の政策の効果について合理性の高い説明を行うことが期待されたのである。それがもしも実現していたなら「説明責任」としての行政機関のアカウンタビリティは大きく向上していたのだろう。

しかしながら実際には、「説明の合理性の向上」は十分に具現化しなかった。各行政機関の側にしてみれば、所管の本務に上乘せして政策効果を綿密に分析するほどの余力もなく、それを分析したところで予算が積み増しされることが保証されるわけでもないという事情が働いていた。もちろん法の命ずるところにより政策効果の追求は意識されていたのかもしれないが、それと同時に政策効果の説明・追求の努力については「評価疲れ」が主張されていたのも事実である。

この論点については、かつてPPBSでの分析手法と同じ種類の壁があった点を指摘しておくべきだろう。行政機関内で説明の合理性水準を向上させようとすればするほど、評価手法は民間企業モデルの目標管理型＝業績測定型評価では済まされなくなる。逆にいえば、説明の合理性水準の向上を諦念したところに業績測定が成立していたはずであったという整理をもっと明確にしておくべきであったのかもしれない。しかし政策評価法は、業績測定型評価に甘んじることなく、プログラム評価も包摂していた。日本の政策評価制度ではこの点についてとくに、評価手法を併記したことで

業績測定型評価とプログラム評価の境界線が曖昧化していた点を指摘しておきたい。それを表現したものが、法第10条における評価書の記載項目の規定にほかならなかったのではないだろうか。

第2に政府システムの改革、とくに分権改革との関係である。政策評価の導入のきっかけとなった中央省庁等改革と同時期には、もうひとつの改革である地方分権改革が展開していた。これらの改革はいずれも権力分立の構造に変容を迫る内容を含むものであった。

1995年から2001年にかけて展開した分権改革は、「第一次分権改革」（西尾勝）とよばれた。その改革のポイントは、「自己決定・自己責任」の原則のもと、国と地方の関係を「上下・主従」から「対等・協力」へと切り替えたこと、機関委任事務や通達を全面的に廃止し、国の指揮・監督権を奪ったこと、国と地方との関係についてルール化（関与のルール化）を行ったことなどであった。これらの改革は国と地方にかかる権力分立のあり方に変容を迫る内容を含むものであった。

国の行政機関のアカウンタビリティについていえば、これから先に展開した一連の分権改革までもを含め、さまざまな課題が噴出した。たとえば、国の関与の縮減は、国の行政機関から指揮監督権を減少させることで国側の説明力の低下をもたらしものとなった。ここでの指摘は分権改革が間違っていたという点を主張しているのではない。国・地方をあわせた日本の政府体系全体の権力分立の再構築に伴い、公共部門全体にかかるアカウンタビリティの構造もまた再構築される必要があったはずであったというのである。

分権改革は端的に言って国と自治体の行政機関同士の権限バランスを変容させるものにほかならなかった。アカウンタビリティ論にとって重要であったのは、権限や予算の移譲とともに、これをコントロールする立場にある政治権力についても、権力の再配置の議論が必要であるという点であった。

この点の議論を十分になしえないまま、国の行政機関のアカウンタビリティを国会と行政機関との関係として従前通りに追求するというこの意味はいかなるものであったのであろうか。このような議論は分権改革論においても希薄であった。

それは権限と責任追及の乖離を招き、アカウントビリティの剥落の可能性を秘めるものであった。付言するなら、新たに権限を付与された自治体の側についても問題があった。行政権限が分権によって拡充されるなか、二元代表・機関対立の関係性もまた従前のままとされたままであった。このことを背景として、議会批判が募り、議会改革が叫ばれるようになったのは周知の通りである。

原理的にいえば、権力分立の修正を図る際には、アカウントビリティの経路の修正もまた必要となる。それが十分に議論されてこなかった点は政府システム改革上の問題点であった。政策評価制度はこのような問題状況の上で展開するという課題を抱えていた。

第3に政策評価法附則第2条の見直しについてである。法施行後3年を経過し、政策評価法附則第2条に基づき政策評価制度の見直しが行われた。そこで掲げられた課題は以下の5点であった（総務省「政策評価制度に関する見直しの方向性」（2005（平成17）年6月17日））。

- 1 評価結果の予算要求等政策への反映
- 2 重要政策に関する評価の徹底
- 3 評価の客観性確保
- 4 国民への説明責任の徹底
- 5 政策評価の基盤整備

ここにはふたたび「国民への説明責任の徹底」の論点が登場した。そこでの議論の内容は、「国民への分かりやすさの確保、評価書・要旨の標準化」と「国民的議論の活性化」であった。後者の内容は、当時各地で行われていた「政策評価フォーラム」やパンフレット、大学での広報活動である。重要なのは前者であった。

前者の「国民への分かりやすさの確保、評価書・要旨の標準化」とは何を意味していたのだろうか。この点については、「一方、各府省においては、実績評価方式を用いた評価が広く用いられているが、評価の観点や実施時期など法定記載事項が必ずしも明確に記載されていない府省もある。」という問題提起がなされていた。この点に対する総務省の処方箋は、「標準的な様式を提示」というものであった。それが「説明責任」とし

てのアカウントビリティとして必要なことと認識されていたのである。ここには政策評価とアカウントビリティをめぐる重要な論点が伏在していた。

一方で行政活動の「説明の合理性の向上」を目指そうとすれば、学術的な手法の摂取、個別政策の個性に寄り添った合理的な分析の展開という方向性をたどることとなる。それは、附則第2条の他の項目、すなわち「評価の客観性確保」が掲げていた論点でもあった。

他方で、説明の合理性水準の向上を諦念ところに成立する業績測定型評価こそ、「標準的な様式」の議論に直結するものであった。それは、国民への説明責任として、評価の総覧性にアクセントをおくところに成り立つものでもあった。評価の総覧性の重視からいえば、個別政策への洞察は犠牲にならざるを得なかった。

「評価の総覧性の重視」と「合理性水準の向上」という二つの説明責任のバランスは、どちらが正しいというものではない。政策評価論では「M（measurement）&E（evaluation）の相互補完性」が強調されていることからそれはうかがえる（山谷 2010）。ただし、この附則第2条の見直しの議論では、それが「府省の政策評価」というひとつの評価の仕組みの上で議論されていた。それは、本来であれば、GAOのプログラム評価とGPRAのように、系統の異なる制度の上で、あるいは異なる機関のミッションの上で追求されるべきものであっただろう。

第4に2009年秋の総選挙で登場した民主党等連立政権での行政刷新会議を中心とする「事業仕分け」「行政事業レビュー」（あわせて「事業仕分け等」）の展開である。このうち行政事業レビューについては2012年12月の総選挙で成立した自民党連立政権においても継承されているところである。

民主党等連立政権は「脱官僚依存・政治主導」を掲げて登場したが、事業仕分け等はその金看板となった。事業仕分け等は、政府レベルと府省レベルの違いはあるものの、外部有識者を交えつつ、政策評価で展開した施策レベルの評価ではなく、「事業レベル」に照準を合わせる点に特徴があった。事業レベルは予算との関連が深く、このため、外部有識者を交えた議論は、実質的な「公開予算査定」の様相を呈し、国民の注目を集めた（南島

2011)。

アカウンタビリティとの関係でいうならば事業仕分け等の取り組みは画期的であった。それは「外部」からの「問責」を、政官関係という権力分立をふまえて組み立てていたからである。国民の選んだ政府が、中立性と継続性とを旨とする行政機関の展開する諸活動に対して「問責」を行うのであるから、その構造からいってもこれは、国民にとっても分かりやすいものであった。

他方で、政策評価が目指していた「合理性水準の向上」からいえば、短時間での審査、結論の明確性、十分な資料や情報が不足するなかでの判断、仕分けの成果を急ぐメディアへの追従などの課題があった。それは、綿密な分析への接近とは逆方向を向いていた。

後者の論点についてさらに、これらの事業仕分け等の議論が財政問題を背景として「効率性」の議論に傾斜していた点もあわせて指摘しておきたい。「合理性水準の向上」は、本来であれば事業効果、すなわち事業の「有効性」をめぐる展開すべき性質のものであった。しかしながら事業仕分け等では、有効性の観点について十分なアクセントがおかれてはいなかった。なぜなら、事業仕分け等は予算編成と絡みながら展開しており、「予算の縮減」を代表的なテーマとしていたからである。いいかえれば予算の議論を行うために事業レベルに注目していたのであって、そうであればこそ「合理性水準の向上」には十分な力点がおかれていなかったのである。そして、「合理性水準の向上」は政策評価の論点として留保され、「政策評価と行政事業レビューとの役割分担と連携のあり方」(総務省)が模索されていった。

様々な課題はあったが、事業仕分け等は政策評価とアカウンタビリティについて重要な一石を投じるものであった。それは「権力分立」と「合理性水準の向上」をめぐる論点であった点をここでは指摘しておきたい。

第5に政策評価における「PDCAサイクルの強調」についてである。先に触れた「標準的な様式」の議論は、行政刷新会議の事業仕分けにおいて、「政策評価、行政評価監視」が「抜本的な機能強化」を言い渡されたのちに本格的に展開することとなった。その指摘を受けての改革の全体像は

2010年の「行政評価等プログラム」(総務省)において示された。

府省の政策評価についていえば、結果としてこの「抜本的機能強化」の延長線上において「標準的な様式」が確立することとなった。府省の政策評価は、先に触れた3つの標準的な評価方式のうち「実績評価方式」を基本とし、それまでの各行政機関の任意的な手法選択というあり方から、総務省が提供する「標準的な様式」に基づいて行われることに改められることとなった。あわせて「事前分析表」の作成(政策体系の整理)も行われることとなり、政策評価の様式統一が一気に押し進められた。総務省が「目標管理型の政策評価」と称していたものがこれである。

「目標管理型の政策評価」のポイントは、「各行政機関における政策体系及び政策のミッションの明確化」「PDCAサイクルを通じたマネジメントの向上」「国民に対する説明責任の徹底」「政策評価と行政事業レビューとの連携の確保、事務負担の軽減」の4点であった。

ここではこのうち「PDCAサイクルを通じたマネジメントの向上」について触れておきたい。それまで「PDCAサイクル」を政府部内で標榜してきたのは経済財政諮問会議や財務省であった(周知の通りPDCAサイクルは民間企業の経営管理において用いられているもので、plan→do→check→actの四段階のサイクルを回すことで経営品質の管理とその向上を目指すというものである)。

ここでは、この「PDCAサイクル」という「管理の論理」が、政策を所管する府省の側の「政策の論理」とは異なる点に注意を促したい。「管理の論理」は組織運営上のトップマネジメントを支えるミドルマネジメントに力点がある。とくに財政逼迫のなかでのミドルマネジメントについては「効率性」を中心とする議論への傾斜が高まりがちである。他方、「政策の論理」は、政策効果の発現をふまえ、その「有効性」に注目し、適切な政策効果が発揮されるように政策体系の見直しを図ろうとするものである。これはすなわち政策を所管する部局の論理である。一方で「管理の論理」を徹底すれば政策を所管する部局の自律性は失われ、他方で「政策の論理」を徹底すればミドルマネジメントや総覧性が喪失しかねない。そうした

葛藤がこれらの間にはあるのである。

ここでは「管理の論理」と「政策の論理」が「効率性」と「有効性」をめぐってときに対立するものであった点を強調しておこう。そして、「目標管理型の政策評価」や「標準的な様式」への舵取りは、「管理の論理」を重視したものである点を確認しておきたい。

政策評価とアカウンタビリティとの関係に戻ろう。ここでの「説明責任」はいったい何を意味しているのだろうか。そこでは繰り返し、行政活動の総覧性が、国民に対する説明責任の徹底方策として示されていた。ここでは「管理の論理」のアカウンタビリティと「政策の論理」のそれとが乖離していた点を確認できるはずである。

4. なぜ、アカウンタビリティは向上しないのか？

政策評価の10年を振り返るとき、本稿がひとつの問いとして掲げていたのは、「この10年の政策評価の経験で国の行政機関のアカウンタビリティは向上したのか」というものであった。それは、これまで見てきたように日本の政策評価制度の重要な目的として掲げられていたものであり、説明責任として繰り返し強調されてきたものであった。

ひとことで先の問いに答えるならば、やはり政策評価制度の導入によっても行政機関のアカウンタビリティは向上したとはいえないということになるだろう。もちろんそれは、アカウンタビリティの理解のあり方如何にもかかわる問題である。

いうまでもなく政策評価を通じて国の行政機関の情報量は圧倒的に増加した。この点において政策評価制度の意義は積極的に評価されなければならない。もしもアカウンタビリティの概念を「管理の論理」のそれとして論じるならば、その限りにおいて政策評価はアカウンタビリティの向上に寄与したということもできよう。

しかし、それは政策評価が本来果たすべき「政策の論理」のアカウンタビリティの確保の議論ではなかった。政策評価によるアカウンタビリティの確保を、「権力分立の確立」「政策実施過程をふまえた議論」「社会科学の諸手法を用いた政策の

調査・分析」「プログラムの観念の重視」「外部評価による中立性・客観性の確保」であると理解するなら、それはこの10年の政策評価の経験によってもさほど進展したとはいえない。すくなくともそれを進展させるような方策は十分に展開してこなかったといっても過言ではない。

(1) アカウンタビリティの確保が進まない理由

この意味において、なぜ日本で政策評価によるアカウンタビリティの確保が進まないのかという理由として指摘しておきたいのは以下の3点である。

第1に「権力分立の確立」「政策実施過程をふまえた議論」「社会科学の諸手法を用いた政策の調査・分析」「プログラムの観念の重視」「外部評価による中立性・客観性の確保」といった要素が、日本の現在の「説明責任」の範疇では語られていない。

アカウンタビリティの歴史的展開や政策評価とアカウンタビリティの広い意味での関係性に着目すると上記のような議論は、当然、アカウンタビリティの射程に含まれるはずである。それは「政策の論理」の議論である。しかし、もしもアカウンタビリティの概念を「管理の論理」のそれへと矮小化するなら、上記のような議論は必ずしも含まれなくともよい。この意味において、わたくしたちはアカウンタビリティの概念を、現在の日本語の理解から離れて、もっと前広に理解していく必要があるだろう。

第2に公的部門における経済合理性の議論が、日本の行政文化にやはりどうしてもなじみにくいという点である。それは行政活動の「法的根拠」や「権限」の問題性である。

日本の行政機関は「法的根拠」や「権限」に基づいて行政活動を展開している。それは日本の行政機関が「合法性」ないし「正統性」といった手続論によって、行政の活動を正当化する行政文化を培ってきたということでもある。「法的根拠」や「権限」によって複雑な議論を決着させるという行政文化においては、経済合理性に基づく説明は軽視されがちである。なぜなら、「いったいなぜその事業を行っているのか」という問いに対しては、「法律があるから」「権限があるから」というのが究極的な回答となってしまうからである。

経済的合理性からいえば、そこには「〇〇のよう効果があるから」という有効性の説明が添えられる必要がある。しかし日本の行政機関においてはこれまでそのような合理性をめぐる説明はあくまでもオブショナルな説明にとどまってきた。もっとも、経済合理性のみに政府活動の合理性を矮小化する点については社会的な批判もすくなくない。だが、政策評価論からしてみれば、それはまた別の議論となる。

第3に、このような議論のすれ違いが、そもそも権力分立絡んでいる点を指摘しておきたい。それは、適正な権力分立が構築されない限り、適正な問責＝答責関係は構築されえないという論点である。

実際のところ、アカウンタビリティが表層的な「説明責任」へと矮小化されてしまうのも、政策評価が分担管理の壁を越えることが難しいのも、いずれもその土台となる権力分立が十分に構築されていない、あるいは現状の権力分立が適正なアカウンタビリティを求める形に整えられていないことを示唆している。これは逆に、厳格な三権分立モデルと評されるアメリカの政治システムの上で、アカウンタビリティと絡む政策評価が発達し得たこととあわせて議論したい論点である。またそれは、政策評価とアカウンタビリティの関係をめぐる歴史的展開の厚みをふまえて振り返っておきたいものでもある。

(2) アカウンタビリティ追求のための政策評価

山谷清志はその著『政策評価』（山谷 2012）において、日本における政策評価論の誤解のポイントとして以下の5点を指摘していた。第1にアカウンタビリティの説明責任への矮小化である。第2にアカウンタビリティの歴史についての理解の不十分さである。第3にプログラムのアカウンタビリティを求めるようになったのは誰かという点についての理解不足である。第4にアカウンタビリティ論についての政策学の貢献と政策評価論への合流についての理解不足である。第5に「評価の結果を誰がどのように使うのか」という点についての誤解である。第6に実務についての理解の不足である。第7に監査と評価の違いについての理解不足である。第8に政策に関する知識の不足で

ある。

これらの山谷の指摘は、行政学の蓄積をふまえていえば、「正しい」といえるものばかりであるだろう。そしてその目指す方向性は、本稿でいうところの「合理性水準の向上」である。もしもこれを追求したいというのであれば、アカウンタビリティ概念を「管理の論理」の方向性だけに発展させようとするだけでは不十分である。もっといえば、企業モデルを念頭におき、マネジメント・サイクルを強調し、業績測定型評価を発展させるだけでは不足しているということになる。

結局のところ、この10年の経験はこのことを示していたように思われる。この10年の政策評価の経験において、政策評価によるアカウンタビリティの追求は「暈の上の水練」ととどまってきたのであり、十分に「実践」へと移行できたわけではなかった。このことは業績測定型評価が抱える課題群が雄弁に物語ってくれている。

政策評価の実践にとって重要なのは、このような意味でいえば、やはりアカウンタビリティとの適切な連結関係の再構築だろう。その際、アカウンタビリティの基礎となる権力分立のあり方にも十分目配せをする必要があるだろう。このうちとくに、政策実施過程の構造面への接近が重要となる。その整理が多少なりともつくのであれば、日本の行政機関の合理性水準はもうすこし改善するのかもしれない。

注記

- 1 アカウンタビリティ論の全体像の整理としては山本（2013）が優れている。その他、西尾（1989）、山谷（1984）、南島（2009）もあわせて参照。
- 2 たとえば、加藤一明・加藤芳太郎・佐藤竺・渡辺保夫（1985）を参照。
- 3 日本における学問分野との関係では、「外在・公式」は憲法学で論じられるコントロール論、「外在・非公式」は政治学（とくに政治過程論）で論じられるコントロール論、「内在・公式」は行政法で論じられるコントロール論、「外在・非公式」は行政学で論じられるコントロール論と整理することもできる。
- 4 行政責任とプロフェッショナリズムについては、山谷（1991）を参照。

5 プログラム評価のアカウンタビリティ・アプローチについては、山谷（1988a）および山谷（1987）に詳しい。

参考文献

- 今村都南雄（2006）『官庁セクショナリズム』、東京大学出版会
- 岩田一彦（2010）「政策評価の『目的』と『情報』」、『評価クォーターリー』13：43-71
- 加藤一明・加藤芳太郎・佐藤竺・渡辺保夫著（1985）『行政学入門：第2版』、有斐閣
- 政策評価研究会編著、行政管理研究センター監修（2008）『詳解・政策評価ガイドブック：法律、基本方針、ガイドラインの総合解説』、ぎょうせい
- 南島和久（2007）「府省における政策評価の中立性および客観性：グリッド／グループ文化理論に基づく考察」、『法学志林』104（4）
- （2009）「行政の信頼性確保と政策評価に関する考察：アカウンタビリティ・システムの再構築」、『季刊行政管理研究』128
- （2011）「府省における政策評価と行政事業レビュー」、『会計検査研究』43：57-71
- 西尾勝（1989）「アカウンタビリティの概念：第1回公会計監査フォーラムの基調講演より」、『会計検査研究』1
- 西尾勝（2001）『新版 行政学』、有斐閣
- 平井文三（2012）「アメリカ連邦政府の職業公務員の人事制度について：GAOを例として」、『亜細亜法學』47（1）
- 益田直子（2010）『アメリカ行政活動検査院：統治機構における評価機能の誕生』、木鐸社
- 山本清（2013）『アカウンタビリティを考える』、NTT出版
- 山谷清志（1984）「行政責任論研究動向に関する一考察」、『大学院研究年報（中央大）』13（1）
- （1987）「合衆国連邦政府における行政統制システムの動向：プログラム評価をめぐって」、『季刊行政管理研究』37
- （1988a）「行政国家における議会統制の可能性：『交換モデル』とプログラム評価」、『法学新法』94：9-10
- （1988b）「議会の復権とその評価：1970年代におけるアメリカ連邦議会の改革をめぐって」、『季刊行政管理研究』43
- （1991）「行政責任論における統制と倫理：学説史として」、『修道法学』13（1）
- 編著（2010）『公共部門の評価と管理』、晃洋書房
- （2012）『政策評価』、ミネルヴァ書房
- C.E. Gilbert（1959）. “The Framework of Administrative Responsibility”, *The Journal of Politics*, 21
- C.J.Friedrich（1940）. “Public Policy and The Nature of Administrative Responsibility”, C.J.Friedrich and E.S.Mason ed., *Public Policy*, Harvard University Press
- C.Hood（1998）. *The Art of the State: Culture, Rhetoric, and Public Management*, Oxford University Press.
- D.Osborne and T.Gaebler（1993）. *Reinventing Government: The Five Strategies for Reinventing Government*, Plume（野村隆・高地高司訳『行政革命』日本能率協会マネジメントセンター、1994年）
- H.Hatry and J.Wholey（1999）. *Performance Measurement: Getting Results*, Urban Institute Press（上野宏・上野真城子訳『政策評価入門：結果重視の業績測定』東洋経済新報社、2004年）
- H.Finer（1940）. “Administrative Responsibility in Democratic Government”, *Public Administration Review*, Autumn（1）

（2013.9.19受理）

Policy Evaluation System and Accountability in Japanese Government

Kazuhisa Najima

Kobegakuin University
najima@law.kobegakuin.ac.jp

Abstract

Ten years have passed since the Government Policy Evaluation Act of Japan was passed. With accountability as one of the purposes of the policy evaluation act, I ask “did the law improve the accountability of Japanese government agencies?”

Results show that while the transparency of government agencies improve through the introduction of the policy evaluation system; the accountability of government agencies did not develop as much. However, the more important issue is why and how the accountability of government agencies did not improve.

The aim of this paper is to clarify, from the standpoint of political science, the relationship between policy evaluation and accountability in government agencies of Japan.

Keywords

Accountability, separation of powers, Government Policy Evaluation Act,
logic of management and policy

【実践・調査報告】

独立行政法人の二次評価に関する考察¹

竹中 一人

前総務省行政評価局 調査官

k2.takenaka@soumu.go.jp

要 約

中央省庁等改革の一環として2001（平成13）年4月1日から独立行政法人制度が発足した。独立行政法人は、国民生活及び社会経済の安定等の公共上の見地から確実に実施されることが必要な事務及び事業であって、国が自ら主体となって実施する必要はないが、民間の主体に委ねると実施されないおそれのあるもの等を実施する法人であり、業務の効率性や質の向上、法人の自律的業務運営の確保、業務の透明性の確保を図る仕組みとなっている。この仕組みと担保するために、独立行政法人が行った事務・事業の業績については、各府省の独立行政法人評価委員会が行う評価とそれら評価委員会が行った評価を再評価する政策評価・独立行政法人評価委員会の二次評価が行われる。

本稿は実践報告として、政策評価・独立行政法人評価委員会が行っている独立行政法人の二次評価に焦点を当て、その取り組みを時系列的に検証する。特にこのうち評価する側、される側から指摘されている「評価疲れ」について考察する。

キーワード

独立行政法人評価、政独委、二次評価、評価疲れ

1. はじめに

独立行政法人制度は10年以上が過ぎ、制度の綻びを修復すべく政府の行政改革推進本部やその下に設けられた行政改革推進会議等²で議論が行われている。また、自民党の2013（平成25）年の参議院選挙公約の中で、独立行政法人改革については、「評価については、評価項目の必要性を精査し、『評価疲れ』を防止するとともに、業務の達成度合いと効率化度合いに分離」等が記述されていた。

総務省に設置されている政策評価・独立行政法

人評価委員会（以下「政独委」）は独立行政法人の事務・事業の業績について二次評価を行う役割を担っている。この二次評価とは、府省に設置された独立行政法人評価委員会（以下「府省評価委」）が当該府省所管の独立行政法人の事務・事業の業績を一次評価したもののメタ・レベルでの評価（メタ評価）するものである。政独委が二次評価を行うにあたっては、府省評価委に評価の着眼として「政策評価・独立行政法人評価委員会における独立行政法人評価に関する運営について」（2002（平成14）年3月22日政独委決定。以下「運営方針」）などを事前に示して評価作業の効率化を図ってき

たところである。この仕組みは制度発足当初から変わっていない。

以下では、政独委の独立行政法人評価分科会（以下「独法分科会」）に課せられた業務³のうち、府省評価委が行った一次評価に対する二次評価の視点等について「独立行政法人制度黎明期」「財務情報の評価の強化期」「最近の独立行政法人評価の傾向」の3つの時代に分けて振り返りながら、制度発足当初からの課題ともなってきた独立行政法人評価における「評価疲れ」について触れることとしたい。

なお、独立行政法人の中期目標の期間の終了時の事務・事業の見直し（以下「勧告の方向性」）については別表に掲げた。

2. 独立行政法人制度黎明期

独立行政法人通則法（以下「通則法」）において、府省評価委は独立行政法人の業務の実績に関する一次評価⁴を行い、政独委は二次評価として府省評価委が行った評価について必要があると認めるときは、意見を述べるができることとなっている⁵。

法令上、政独委と府省評価委とは対等な立場であり、府省評価委はそれぞれ独自の評価基準を持っている。政独委は、運営方針において、二次評価の実施方法として「当委員会（政独委）の評価においては、まず、各府省評価委員会の評価結果が当該評価委員会において定められた評価基準に適合したかたちで適切に評価を行ったものとなっているか、また、評価基準を踏まえた評価の内容は妥当なものとなっているかについて点検を行うことを基本とする」ことや、政独委の立ち位置を「政府全体レベルの評価機関である当委員会」と表明し、府省評価委が行う一次評価と政独委が行う二次評価の役割を明確化し、二次評価はメタ評価であることを宣言していた。政独委は独立行政法人の業務の実績の二次評価を行う立場から、どのような視点で二次評価を行うかという方針を政独委決定として予め示したのである。

（1）最初の二次評価

具体的に運営方針においては、「全政府レベルの評価機関である当委員会は、評価の一環として、独立行政法人評価を通じた実効性向上を図る見地から、各府省評価委員会の評価結果を全体的、横断的に把握し、評価の実効性向上に資すると考えられる一定の手法、視点等が・・・評価の実効性効能のため有効かつ必要と認められる場合には、当委員会から各府省評価委員会に対して通知等を行うこととし、また、次年度以降の評価結果の評価において、当該手法、視点等の検討状況、活用状況について把握を行うこととする。」と記述されていた。

その後、政独委は「平成13年度における独立行政法人の業務の実績に関する評価の結果についての第1次意見」（2002（平成14）年11月19日政独委決定。以下「1次意見」）、「平成13年度における独立行政法人の業務の実績に関する評価の結果についての第2次意見～独立行政法人評価の厳格性・信頼性の向上のために～」（2002（平成14）年12月26日政独委決定。以下「2次意見」）を公表した。

1次意見は、府省評価委の評価結果が、当該評価委において定められた評価基準に適合したかたちで適切に評価を行ったものとなっているか、また、評価基準を踏まえた評価の内容は妥当なものとなっているかについて点検を行い、その結果を政独委から当該評価委に通知したものであった。これは、現在もその通知スタイルを踏襲しているものである。

2次意見は、同文書の前文において、「各府省の独立行政法人評価委員会がそれぞれの基準に基づいて評価を行うことのメリットを確保しつつ、政府全体として、評価の厳格性・信頼性を向上させるため重要な事項について、適切な評価活動を確保することを狙いとしており、いわば共通の『評価活動準則』に当たるものとして機能することが期待される。当委員会は、各府省の独立行政法人評価委員会において本意見の具体化に向け積極的かつ適切な対応が行われ、次回以降の評価において逐次反映されることを要望する」と記述しており、府省評価委の評価基準の共通化を図ろうとしたものであった。

(2) 2回目の二次評価

政独委は独法分科会の下で、「財務内容の改善等についての評価方法の在り方に関する研究会」を設置した。同研究会では、府省評価委が行う平成14年度における業務の実績に関する評価のうち、「予算、収支計画及び資金計画の実施状況等」および「業務運営の効率化に関する目標を達成すべき措置の実施状況」に関する事項を、政独委として特に関心を持つべきものとして整理した。これが、2003（平成15）年7月31日に開催された政独委独法分科会において「平成14年度業務実施評価の結果についての評価における関心事項（「財務内容の改善」および「業務運営の効率化」関係）」として報告され、当該報告と2次意見とを併せた評価視点が政独委の行う二次評価の視点とされた。

一方、2002（平成14）年度にはじめて二次評価を受けた府省評価委の反応について、政独委が府省評価委に通知した「平成14年度における独立行政法人の業務の実績に関する評価の結果についての意見について」（2003（平成15）年11月13日政独委決定）において評価自体が法人に過度の負担になっていること、そのために独立行政法人の自主性を損なうことになるとの懸念を記述していた⁶。これらの懸念等を踏まえ、政独委は「今回の評価においては、評価の在り方の観点をも勘案しつつ評価作業に取り組んでまいりました」としながらも、①評価のために把握するデータは、評価のためだけに把握・分析が求められるものではなく、業務運営上の自律性を与えられた法人の長が法人運営の判断を行う際に把握・分析すべきものであること、②独立行政法人の特殊性から、自主的な運営の結果を業務面、財政面全般にわたって国民にわかりやすい形で明らかにし、評価を受ける等高度な透明性を確保しなければならない、との認識に立ち、2002（平成14）年度の2次評価を行ったことを表明した⁷。

(3) 3回目の二次評価

国民生活センター等34の特殊法人等から独立行政法人に移行してきた法人（以下「移行独法」）の評価をはじめて行う年度には、政独委としても二次評価を行う前から周到な準備を積み上げていた。具体的には、2004（平成16）年3月以降、独

法分科会内に、①研究開発関係法人の評価方法の在り方に関する研究会、②教育・指導・訓練関係法人の評価方法の在り方に関する研究会、③振興助成・融資関係法人の評価方法の在り方に関する研究会、④公共用物・施設設置運営関係法人の評価方法の在り方に関する研究会、⑤財務内容の改善等についての評価方法の在り方に関する研究会を設置し、法人の業務類型等に応じて求められる見直しの視点や二次評価の観点について検討を行っていた。

2004（平成16）年6月30日に独法分科会においては、「独立行政法人の見直しの前倒しと特殊法人等移行法人の厳格な評価への対応～委員会審議の前倒しと『評価における関心事項』の取りまとめ～」を発表し、移行独法の評価について慎重が期された⁸。また、同日付で「平成15年度業務実績評価の結果についての評価における関心事項」も発表し、移行独法における特殊法人時代の資産の評価の在り方について特段の関心があることが表明された⁹。

政独委が府省評価委に通知した「平成15年度における独立行政法人の業務の実績に関する評価の結果についての意見について」（2004（平成16）年12月10日政独委決定）では、政独委の二次評価として「評価における関心事項」を踏まえ、二次的、横断的な評価を行い必要な意見を取りまとめたことや、2005（平成17）年度に中期目標終了時の見直しが予定されている法人については、勧告の方向性につながる主要な論点について府省評価委の評価が的確に行われるよう促す観点から意見が取りまとめられた¹⁰。

3. 財務情報の評価の強化期

2005（平成17）年度は独立行政法人の第1期の中期目標の期間が終わり、第2期中期目標期間に入る法人が多くなってきたことで、政独委や府省評価委の評価作業も落ち着いてきたかに見えた時期である。しかし、移行独法では財務状況が不明瞭なものもあり、不要財産の処分や契約の適正化などが注目されていた。そのため、政独委としても財務諸表等やセグメント情報など独立行政法

人の財務状況について重点的に評価を行うようになっていった。

(1) 4回目の二次評価

政独委は「平成16年度業務実績評価の結果についての評価における関心事項（「財務内容の改善」及び「業務運営の効率化」関係）」（2005（平成17）年7月11日政独委決定）を府省評価委に示した。この「評価における関心事項」は、その後の二次評価の在り方を示す転換点となったものであり、それまでの「各府省評価委員会の評価結果を全体的、横断的に把握」する方針から、将来の勧告の方向性や政独委が横断的に把握・「分析した結果」を踏まえたものに重点化する方向へと舵を切り始める契機となった文書である¹¹。

この「評価における関心事項」では、独立行政法人の財務諸表等の財務情報を業務実績評価等に活用していくことについて、①運営費交付金の収益化について多くの法人が費用進行基準を導入していない、②中期目標・中期計画の業務と財務諸表等の財務情報が連動していない等の課題が指摘されていた。

政独委が府省評価委に通知した「平成16年度における独立行政法人の業務の実績に関する評価の結果についての意見について」（2005（平成17）年11月14日政独委決定）において、「各府省評価委委員長との懇談会における意見等も考慮し、中期的観点をも踏まえた業務の在り方の検討に資する指摘、効果的・効率的な法人業務運営の観点からの主要な事項にかかる指摘等に重点を置いて」、二次的、横断的な評価を行い、必要な意見を取りまとめたことが表明されていた。

あわせて、府省評価が行った一次評価で参考となる手法等の紹介、評価の取り組みが顕著な府省評価委や積極的な取り組み姿勢が顕著な独立行政法人の例について参考通知することとなった¹²。

(2) 5回目の二次評価

2006（平成18）年度末および2007（平成19）年度末には、中期目標期間が終了する法人が40法人と多数であった。これらの中にははじめて中期目標期間が終了する法人が含まれていたことから、政独委には勧告の方向性を円滑かつ効果的に行う

必要性があった。このため、「平成17年度における独立行政法人等の業務の実績に関する評価の結果等について意見について」（2006（平成18）年11月27日政独委決定。以下「18年度方針」）が例年よりも約半年早く各府省評価委に示された¹³。

18年度方針の内容は、独立行政法人の実施する業務の質の確保を図りつつ、業務運営の効率性、自主性を高めるとともに、国の歳出の削減を図る観点から、①業務の廃止・縮小・重点化、②経費の縮減、業務運営の効率化、③自己収入の増加、④情報提供（ディスクロージャー）の充実の4つの視点を基本とするものであった。また、「評価における関心事項」における業務類型ごとの視点なども盛り込まれていた。さらに、『平成18年度における独立行政法人の組織・業務全般の見直し方針』（2006（平成18）年7月18日政独委決定）における視点等も考慮し、必要な意見を取りまとめたことが触れられていた。

(3) 6回目の二次評価

2007（平成19）年4月には独立行政法人緑資源機構の官製談合事件が発覚した¹⁴。同事件を契機とした閣議決定¹⁵を踏まえて、政独委は「独立行政法人の中期目標期間終了時の見直し及び業務実績評価に関する当面の取組方針」（2007（平成19）年7月11日政独委決定。以下「業務実績評価に関する当面の方針」）を示した。その中で「一部の独立行政法人の業務に関する不祥事案が相次いで明るみになっており、遺憾ながら、個別の独立行政法人に対する国民の信頼が失われるだけでなく、制度全体が不振の目で見られ、その根幹が揺らぎかねない事態になりつつある」「制度に対する国民の信頼回復につながるような厳しい取組が不可欠であり、当委員会としても各法人の事務・事業について、聖域を設けることなく、これまで以上に厳しい態度で見直しを行う」「業務実績評価に関し、各府省の独立行政評価委員会における取組が不十分ではないかとの指摘が国会での議論やマスコミ報道等でなされていることも事実である。各府省の独立行政法人評価委員会においては、この状況を虚心坦懐に受け止め、従来の評価の客観性・厳格性について検証した上で、評価の質の更なる向上のため研さんを積むことが求められて

いる」等と記載された。独立行政法人制度の柱である評価制度に対して危機感が表明されたのである¹⁶。

(4) 7回目の二次評価

「平成19年度業務実績評価の取組について」（2008（平成20）年7月14日政独委決定）を公表し、業務実績評価に関する当面の方針に基づいて評価することを府省評価委に示した。その中で、特に着目する事項として、欠損金・余剰金・不良債権・交付金債務残高、既往の政独委指摘事項等の6項目が示された。さらに政独委では、「入札・契約の適正化に係る評価における関心事項」（2008（平成20）年9月20日政独委独法分科会随意契約等評価臨時検討チーム）を示し、監事による契約の合規性に係るチェックプロセス等について具体的に評価することとされた。

その後、政独委委員長は、「平成19年度における独立行政法人等の業務の実績に関する評価の結果等について意見について」と題する談話を発表するとともに、政独委は府省評価委に二次評価の結果（2008（平成20）年11月26日政独委決定）を通知した。委員長談話においては、各法人共通事に6事項および個別法人について56事項を指摘したことが述べられていた¹⁷。

4. 最近の独立行政法人評価の傾向（8～10回目の二次評価と独法評価の視点）

以上の7回にわたる政独委の二次評価の実績を踏まえて、「独立行政法人の業務の実績に関する評価の視点」（2009（平成21）年3月30日政独委決定。以下「独法評価の視点」）を策定した。以降、この「独法評価の視点」に基づいて独法分科会が毎年度作成する「〇年度業務実績評価の具体的取組について」によって、政独委として年度ごとに関心のある評価事項が府省評価委に予め示されることとなり、評価の効率化が取り組まれることとなった。

「独法評価の視点」では、独立行政法人の業務の実績の評価において共通的に求められる基本的な視点として、①法人の業務に係る政策目的を踏ま

えて、その業績を評価していること、②評価に際しては、常に、効率性、生産性等の向上による業績の増進、業務の対象となる国民に対するサービスの質の向上を志向していること、③法人の業務の内容、業績の分析とそれに基づく評価、課題と展望を国民に分かりやすく説明することにより、法人業務に対する国民の理解を深めることを志向していることの3点が列挙されている。政独委はこれらの事項に問題意識を持ちながら、府省評価委の評価結果の適正性が確保されているかについて評価することとなった。

また、「独法評価の視点」においては、各法人に共通する個別的な視点として、①政府方針等、②財務状況、③保有資産の管理・運用等、④人件費管理、⑤契約、⑥内部統制、⑦関連法人、⑧中期目標終了時の見直しを前提とした評価、⑨業務改善のための役職員のイニシアティブ等についての評価の9事項が示されていた。

さらに、総務省は、会社法（2005（平成17）年法律第86号）および金融商品取引法（1949（昭和23）年法律第25号）において、法の目的に応じて必要な内部統制制度が民間企業に導入されたことや独立行政法人整理合理化計画（2007（平成19）年12月24日閣議決定）等を踏まえ、2009（平成21）年7月24日以降、9回にわたり「独立行政法人における内部統制と評価に関する研究会」（以下「内部統制研究会」）を開催し、独立行政法人の内部統制の考え方を整理した「独立行政法人における内部統制と評価について」（2010（平成22）年3月23日）を発表し、「独法評価の視点」を改正（2010（平成22）年5月31日政独委決定）した。

政独委は、「独法評価の視点」を踏まえてより効率的・効果的に二次評価を行うこととしている。ここでは、8回目（2008（平成20）年度）から11回目（2012（平成24）年度）の政独委の二次評価について、特に特徴的なものを整理する。

(1) 8回目の二次評価

2008（平成20）年度の政独委の二次評価に併せて、政独委は独立行政法人ごとの契約状況、給与の諸手当、福利厚生状況について実態調査を行った結果を府省評価委に通知した。

(2) 9回目の二次評価

独立行政法人の不要資産の国庫納付等について定めた独法通則法の一部改正法が2010（平成22）年5月に公布されたことや内部統制研究会の報告等を踏まえて、政独委は「21年度業務実績評価の具体的取組について」（2010（平成22）年5月31日独法分科会決定）を策定した。

(3) 10回目の二次評価

2011（平成23）年3月11日に東日本大震災が発生したため、政独委は独立行政法人が被災者支援や復旧・復興対策を行っていることを踏まえ、被評価者である独立行政法人や評価者の負担等に配慮し、政独委の二次評価はより効率的・効果的に行う観点から、独法評価の視点の「第2 各法人に共通する個別的な視点」から3事項（保有資産の管理・運営等、内部統制、政独委が主務大臣に通知した勧告の方向性のうち平成22年度において取り組むこととされている事項についての法人の取組状況）を抽出し、政独委の関心事項を予め絞って「22年度具体的取組方針」（2011（平成23）年4月26日独法分科会決定）¹⁸を策定した。

(4) 11回目の二次評価

「23年度具体的取組方針」（2012（平成24）年5月21日独法分科会決定）についても、2010（平成22）年度具体的取組方針の方向性を踏襲し、政独委は独法評価の視点のうち、「第1 基本的な視点」関係と、「第2 各法人に共通する個別的な視点」関係から4事項（政府方針等、保有資産の管理・運用等、内部統制、業務改善のための役職員のイニシアティブについての評価）を抽出して評価の重点化を行った¹⁹。

5. 独立行政法人評価の二次評価の課題の解決に向けて

(1) 最新の政独委の二次評価の取組

政独委が行ってきた独立行政法人の二次評価の実際の概要に触れてきたが、その中で、政独委は評価項目の重点化や評価作業の効率化を行おうと何度も試みられてきた。しかし、独立行政法人自

体の増加²⁰、法人の不祥事、無駄遣いの指摘²¹などがあり、それらの事象を真摯に受け止め独立行政法人の評価を行ってきたが故に、政独委は「重点化・効率化」の試みが必ずしも満足になしえていないと料する。

この課題を解消するための手段の一つとして、2013（平成25）年5月20日、独法分科会は二次評価の評価項目の重点化や評価作業の効率化を目指した「平成24年度具体的取組」を決定した。この決定の特徴は、政独委が行う二次評価の内容を重点事項と点検事項とに分け、重点事項については今まで政独委が指摘してきた内容等について重点的にチェックするが、点検事項については今まで二次評価を行って培ってきたノウハウを最大限活用して評価漏れがないか等形式的なチェックを中心とすることで、コストパフォーマンスの高い評価作業を行う方針を示した点にある。

(2) 二次評価の作業の重複感・負担感の軽減対策に向けて

従来から政独委は二次評価の取組方針として「重点化」「効率化」を掲げてきた。独立行政法人の事務・事業は、国や特殊法人等の時代に比べて、活動状況、財務状況、サービス提供に関するコスト削減などの情報を幅広く公開したため法人運営の透明化が進んだ。これらの取り組みには府省評価委や政独委の評価活動は大きく貢献している。一方で、府省評価委、政独委だけでなく評価を受ける独立行政法人にも評価作業に対する負担感・重複感を生じてきたようである。この原因としては業務の体験上以下のことが考えられる。

そもそも政独委の二次評価は府省評価委のお手盛り防止の機能を担ってきた。政独委はダブルチェックを前提としてきた機関である。このため政独委はある程度、評価作業に重複感が出ることを容認されている。しかし、重複感が出ることを前提としても、府省評価委から見れば、政独委は一次評価に対して、必要があると認めるときは意見を述べることで「ができる」²²だけであった。府省評価委は専門機関として十分な評価をしているという自負があるため、政独委からの意見を受けてもう一度評価をやり直すことになると、府省評価委は不本意であろうし、評価を受け直す独立行政

法人自体も再度評価を行うための資料作りなどが発生することとなり重複感は免れない。

また、制度上には現れないが、当然、独立行政法人自体も自らが実施した事務・業務に対する評価を行っているため、独立行政法人自体の評価、府省評価委の評価、政独委の評価で同じ事務・事業についてそれぞれの評価機関が違う角度から意見を言えば、評価を受ける側の負担感・重複感は相当なものとなるだろう。

さらに、事務・事業の評価については、評価書の様式や評価方法・評語など評価の基本となる事項が府省評価委間で異なっているために、政独委は府省評価委間の評価方法等を比較する点については難しさがある。また、独立行政法人XとYで実施された補助金の認定業務などについて、Xの処理期間の目標処理日数が数値化されたもので評価していればYでも同じ評価方法ができないか政独委が指摘しても、Yの業務改善には寄与するが、XとYの業務は異なるため処理日数について比較し評価することは困難である。また、独立行政法人にとってみれば、煩雑な作業が一つ新たに増えたと捉えるだろう。政独委は制度発足時当初から評価をする側・受ける側の負担感・重複感をできるだけ軽減するため、1次意見および2次意見等が「評価活動準則」になることを期待していた。しかし、制度上の拘束力がなく府省評価委の参考資料にとどまってきた。

2013年現在、独立行政法人、府省評価委、政独委にまたがる評価作業への負担感・重複感を排除する等の制度改革の検討が行われているところである。2013（平成25）年6月5日に開催された政府の行政改革推進会議では、寺田内閣府副大臣から、目標・評価の仕組みの見直しとして「主務大臣の下での政策のPDCAサイクルを強化する観点から、主務大臣から法人に対し的確かつ明確な目標を付与することとするとともに、主務大臣自らが業績評価を行う。その際、主務大臣による評価等の客観性や政府全体としての整合性の確保等のため、外部から点検する仕組みを導入するとともに、政策評価と独立行政法人評価の連携を図る。」とした独立行政法人改革に関する有識者懇談会の中間とりまとめが紹介され、制度改革の方向性が示された。主務大臣が行った独立行政法人の事務・

事業の評価を外部から点検できるようにするためには、共通の目標設定方法、評価書の書き方、評価基準、評語等が設定されなければならない。これらの検討も内閣官房等関係府省においてすでに始まっているところである。

注記

- 1 本稿の事実関係やその内容は筆者の実務上の経験に基づくものである。ただし、文中の意見に関する部分や分析については筆者の個人的なものである。
- 2 例えば、2013（平成25）年6月5日に開催された政府の行政改革推進会議では、寺田内閣府副大臣から「（独立行政法人の）目標・評価について、主務大臣は目標を示すのみで評価に関与しないこと、主務大臣の目標が不明確であること等により適切なPDCAサイクルが確立できていない。また、各法人に対し多層的な評価が行われており、主務府省、法人ともに負担が増加している」旨の発言があった。
- 3 通則法における業務として、各府省の独立行政法人評価委員会が行った業績評価に対する意見を述べることおよび中期目標期間終了時に主要な事務・事業の改廃に関して主務大臣に勧告を行うことがある。その他、「独立行政法人、特殊法人及び認可法人の役員の退職金について」（2003（平成15）年12月9日閣議決定）に基づき、独立行政法人の役員の退職金の業績勘案率について、各府省評価委に意見を述べることができるとされている。
- 4 通則法第12条第2項、第32条第1項、第34条第1項等。
- 5 通則法第32条第5項、第33条第3項等。
- 6 「平成14年度における独立行政法人の業務の実績に関する評価の結果についての意見について」の政独委委員長談話の中で「当委員会の昨年の意見（2次意見）については、各方面からいただいたコメントの中に、当委員会の意見に沿ってデータを把握し、評価を行うと、評価自体が法人に過度の負担をもたらすのではないかと、また、独立行政法人の業務運営の自主性を損なうことになるのではないかと懸念する声があった」と記載されている。
- 7 「平成14年度における独立行政法人の業務の実績に関する評価の結果についての意見について」の政独委委員長談話。なお、2003（平成15）年11月13日の政独委では、文部科学省所管の教員研修センターに

- ついて初めて主要な事務および事業の改廃に関する勧告の方向性が決定したことでも注目を集めた。
- 8 「独立行政法人の見直しの前倒しと特殊法人等移行法人の厳格な評価への対応～委員会審議の前倒しと『評価における関心事項』の取りまとめ～」における政独委独法分科会長談話では「『評価における関心事項』は、独立行政法人の主要な事務・事務業を研究開発関係、教育・指導・訓練関係、公共用物・施設設置運営関係、振興助成・融資関係等に分類して、それぞれ固有の評価の視点を取りまとめたものであり、同時に、財務内容の評価の視点も示した。・・・今回発表する『評価における関心事項』に沿って、・・・、特殊法人等から移行した法人の厳格な二次評価を行ってまいりたい。」と記載されている。
- 9 「平成15年度業務実績評価の結果についての評価における関心事項」において「特殊法人等時代の資産が資産評価委員会の時価評価を経て包括継承されていることなどから、評価に際しては特殊法人等時代に比べて財務内容の改善がなされたかの観点や、連結対象となる特定関連会社及び関連企業を有する法人の評価について連結グループ全体の資金の流れやグループ全体のマネジメントシステムを把握する必要があること」が記載されている。
- 10 なお、2003（平成15）年度末をもって中期目標期間が終了した教員研修センターについて、最初の中期目標に係る業務の実績に関する評価が行われた。本件に対する政独委の二次評価は、2003（平成15）年度に行った勧告の方向性や文部科学省評価委の評価結果が活用して新中期計画の業務を励行するものとなっている。
- 11 「平成16年度業務実績評価の結果についての評価における関心事項（『財務内容の改善』および『業務運営の効率化』関係）」において「政策評価・独立行政法人評価委員会が行う二次評価については、評価活動も4年目を迎え、各府省の独立行政評価委員会も評価の蓄積があること、政策評価・独立行政法人評価委員会の評価活動も昨年度と同様に、事務・事業の見直し業務と並行して行われることから、年度意見として述べるものは、将来の事務・事業の見直し（勧告の方向性）を視野に入れたものや横断的に把握・分析した結果を踏まえたものに、重点化する必要がある。」と記載されている。
- 12 政独委は、通則法の評価制度を準用している国立大学法人および大学共同利用機関法人について、それらの業務の実績に関する評価の結果についての意見を国立大学法人評価委員会に初めて述べた。
- 13 その他2006（平成18）年度方針を早く示した背景には、「政策金融改革の基本方針」（2005（平成17）年11月29日経済財政諮問会議）、「行政改革の重要方針」（2005（平成17）年12月24日閣議決定）、「18年度以降当面の独立行政法人の見直しの基本方針について」（2006（平成18）年5月23日行政減量・効率化有識者会議）、「経済財政運営と構造改革に関する基本方針2006」（2006（平成18）年7月7日閣議決定）など行政改革・予算編成の動きと深く関わっていたためである。
- 14 独立行政法人整理合理化計画（2007（平成19）年12月24日閣議決定）で同機構の廃止が決まり、その業務の一部は森林総合研究所および国際農林水産業研究センターが承継することになった。
- 15 「経済財政改革の基本方針2007（2007（平成19）年6月19日閣議決定）」において、「現行の独立行政法人が制度本来の目的にかなっているか等について、原点に立ち返って見直す」こととなり、「独立行政法人整理合理化計画」（2007（平成19）年12月24日）が策定された。
- 16 「平成18年度における独立行政法人等の業務の実績に関する評価の結果等についての意見」（2008（平成20）年1月政独委決定）においても、政独委は府省評価委に対して「一部の独立行政法人における不祥事の発覚等により独立行政法人全体が不信の目で見られていること、業務実績評価に関する各府省の評価委員会における取組が不十分ではないかとの指摘があること」を踏まえて、厳格な評価を行うよう求められた。また、独立行政法人整理合理化計画の策定を踏まえ、評価の更なる質の向上が求められた。なお、通則法の評価制度を準用している日本司法支援センターについて、その業務の実績に関する評価の結果についての政独委としての意見を日本司法支援センター評価委員会に対して初めて表明したのもこの時期のことである。
- 17 その他、同談話では、独立行政法人の給与水準および総人件費改革の取り組みについて、横断的な分析結果を取りまとめたこと等が記述されている。なお、2008（平成20）年度末に中期目標期間が終了する独

立行政法人はなかった。

- 18 また、「平成22年度における独立行政法人等の業務の実績に関する評価の結果等についての意見」(2011(平成23)年12月9日政独委決定)では、東日本大震災による各独立行政法人の業務への影響および震災対策関係の業務の実施状況が取りまとめられ、公表された。
- 19 また、「平成23年度における独立行政法人等の業務の実績に関する評価の結果等についての意見」(2013(平成25)年1月21日政独委決定)では、独立行政法人の自然災害等に関するリスクへの対応および自然災害等の外部要因に起因するリスクに対する法人の自発的取組状況について、それぞれ取りまとめられ、公表された。
- 20 独立行政法人の現況については、2001(平成13)年4月1日に57法人が独立行政法人としてスタートしたが、2013(平成25)年4月1日現在、101法人(公務員型8法人、非公務員型93法人)となっている。
- 21 例えば、「独立行政法人改革の検討の視点」(2013(平成25)年2月27日、第1回行政改革推進会議資料)参照。
- 22 独法通則法第32条第5項。

参考文献

行政改革会議(1997)『最終報告』

政策評価独立行政法人評価委員会(2002a)『政策評価・独立行政法人評価委員会における独立行政法人評価に関する運営について』

—————(2002b)『平成13年度における独立行政法人の業務の実績に関する評価の結果についての第1次意見』

—————(2002c)『平成13年度における独立行政法人の業務の実績に関する評価の結果についての第2次意見～独立行政法人評価の厳格性・信頼性の向上のために～』

—————(2003a)『平成14年度業務実施評価の結果についての評価における関心事項(「財務内容の改善」及び「業務運営の効率化」関係)』

—————(2003b)『平成14年度における独立行政法人の業務の実績に関する評価の結果についての意見について』

—————(2004a)『独立行政法人の見直しの

前倒しと特殊法人等移行法人の厳格な評価への対応～委員会審議の前倒しと「評価における関心事項」の取りまとめ～』

—————(2004b)『平成15年度業務実績評価の結果についての評価における関心事項』

—————(2004c)『平成15年度における独立行政法人の業務の実績に関する評価の結果についての意見について』

—————(2005a)『平成16年度業務実績評価の結果についての評価における関心事項(「財務内容の改善」及び「業務運営の効率化」関係)』

—————(2005b)『平成16年度における独立行政法人の業務の実績に関する評価の結果についての意見について』

—————(2006a)『平成18年度における独立行政法人の組織・業務全般の見直し方針』

—————(2006b)『平成17年度における独立行政法人等の業務の実績に関する評価の結果等について意見について』

—————(2007a)『独立行政法人の中期目標期間終了時の見直し及び業務実績評価に関する当面の取組方針』

—————(2007b)『平成19年度における独立行政法人等の業務の実績に関する評価の結果等について意見について』

—————(2008a)『独立行政法人の業務の実績に関する評価の視点』

—————(2008b)『平成20年度における独立行政法人等の業務の実績に関する評価の結果等について意見について』

—————(2009)『平成21年度における独立行政法人等の業務の実績に関する評価の結果等について意見について』

—————(2010)『平成22年度における独立行政法人等の業務の実績に関する評価の結果等について意見について』

—————(2011)『平成22年度における独立行政法人等の業務の実績に関する評価の結果等について意見について』

—————(2012)『平成23年度における独立行政法人等の業務の実績に関する評価の結果等について意見について』

政策評価・独立行政法人評価委員会独立行政法人分科

会（2007）『平成19年度業務実績評価の取組について』

—————（2008）『平成20年度業務実績評価の
具体的取組について』

—————（2009a）『平成21年度業務実績評価
の具体的取組について』

—————（2009b）『平成22年度業務実績評価
の具体的取組について』

—————（2010）『平成22年度業務実績評価の
具体的取組について』

—————（2011）『平成23年度業務実績評価の
具体的取組について』

—————（2012）『平成24年度業務実績評価の

具体的取組について』

政策評価・独立行政法人評価委員会独立行政法人分科
会随意契約等評価臨時検討チーム（2007）『入札・
契約の適正化に係る評価における関心事項』

独立行政法人制度研究会編（2004）『改訂 独立行政
法人制度の解説』、第一法規

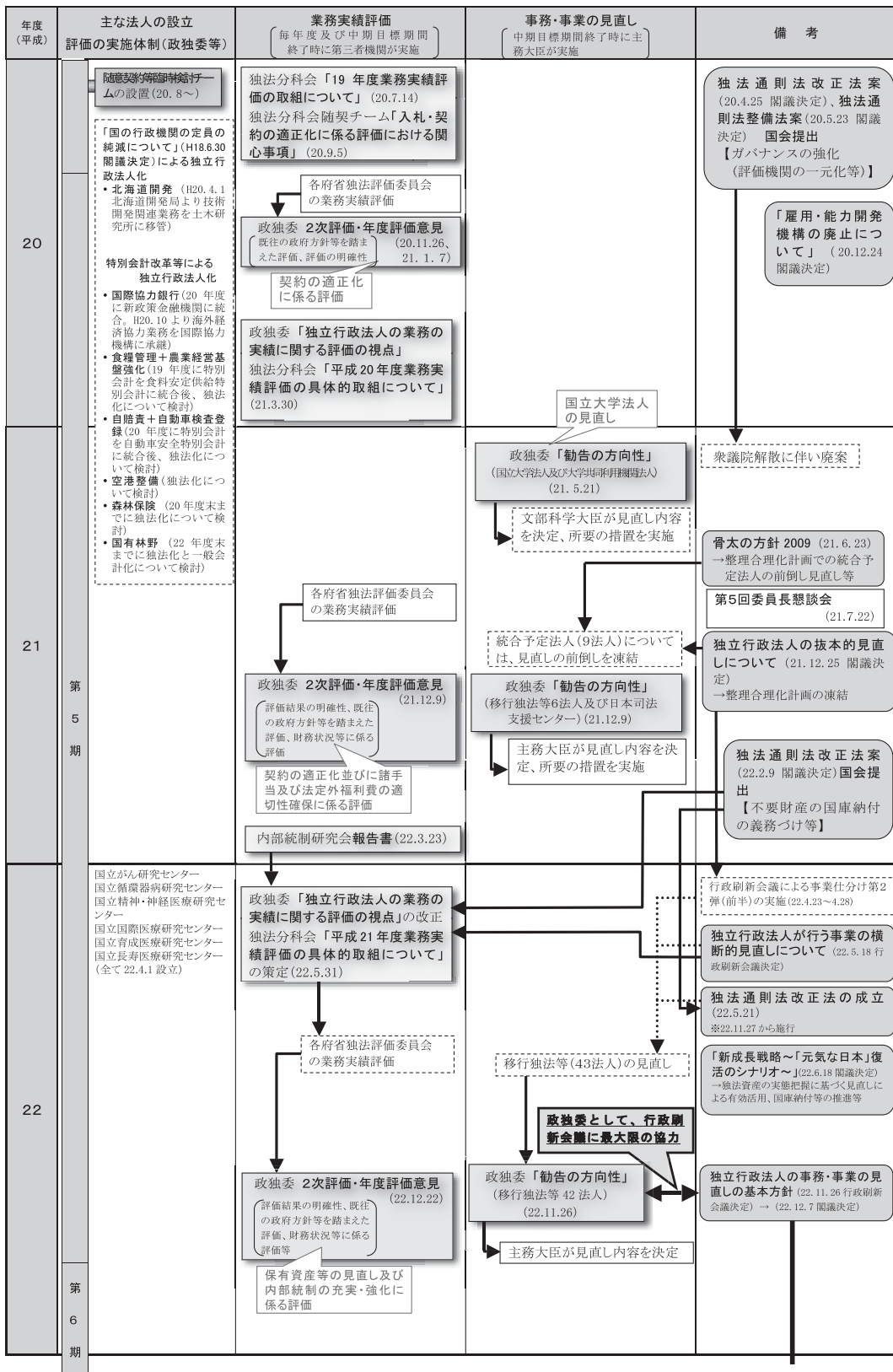
独立行政法人における内部統制と評価に関する研究会
（2008）『独立行政法人における内部統制と評価につ
いて』

宮内忍（2004）「主務省の独立行政法人評価委員会に
よる業務実績評価等について」、『公会計委員会研究
報告』10号

（2013.9.2受理）

表 独立行政法人評価の主な経緯

年度 (平成)	主な法人の設立 評価の実施体制(政独委等)	業務実績評価 〔毎年度及び中期目標期間 終了時に第三者機関が実施〕	事務・事業の見直し 〔中期目標期間終了時に主 務大臣が実施〕	備 考
12	政策評価・独立行政法人 評価委員会(政独 委)が発足 (13.1.23)	独立行政法人制度がスタート (13.1.6)		
13	先行独法 57 法人 設立 (13.4.1)	政独委「独立行政法人評価に 関する運営について」(14.3.22)		「特殊法人等整理合理化計 画」(13.12.19 閣議決定) →特殊法人等の独法化
14	第 1 期 府省別に3ワーキ ング・グループ(W G)を開催 (14.9～) 財務研究会を開催 (14.12～)	各府省独法評価委員会 の業務実績評価 政独委 2次評価・年度評価意見 (1次意見・個別事項) (14.11.19) 政独委 年度評価意見 (2次意見・共通事項) (14.12.26)	最初の年度 評価意見	総理と独法評価委員会委員 長との懇談 (14.7.9) 第1回委員長懇談会 (14.8.30)
15	第 2 期 移行独法 30 法人 設立 (15.10.1) 業務類型ごとに4 つの研究会を開催 (16.3～)	財務研「評価における関心事項」 (15.7.31) 各府省独法評価委員会 の業務実績評価 政独委 2次評価・年度評価意見 (15.11.13)	政独委「勧告の取組方針」 (15.7.1) 「中期目標期間終了時における独立行政 法人の組織・業務全般の見直しにつ いて」(15.8.1 閣議決定) 政独委「勧告の方向性」 (教員研修センター) (15.11.13) 主務大臣が見直し内容を決定、 所要の措置を実施	第2回委員長懇談会 (14.5.29) 「勧告の方向性」 を規定 最初の見直し (先行独法)
16	国立大学法人設 立 (16.4.1) 中期目標期間は平成 21年度までの6年間 WGを5つに拡大 (16.8～)	4研究会・財務研「評価における関心事項」 (16.6.30) 〔①研究開発関係法人、②教育・指導・訓練関係法人、③公共用物・施設設置・ 運営関係法人、④振興助成・融資関係法人、⑤財務内容の改善等〕 各府省独法評価委員会 の業務実績評価 政独委 2次評価・年度評価意見 〔事務・事業の見直しにつな がる評価 等〕 (16.12.10) 移行独法の 年度評価	政独委「勧告の方向性」 〔消防研究所など 先行独法 32 法人〕 (16.12.10) 主務大臣が見直し内容を決定、 所要の措置を実施	骨太の方針 2004 (16.6.4) →前倒し見直し等を規定 行政改革推進本部に独立行政 法人に関する有識者会議を設置 (16.6.17) 有識者会議指摘事項 (16.10.27)
17	第 3 期 日本高速道路保有・債務 返済機構 (17.10.1) 年金・健康保険福祉施設 整理機構 (17.10.1) 政策金融WGを開催 (18.2～)	財務研「評価における関心事項」 移行独法・財務評価の視点 (17.7.11) 各府省独法評価委員会 の業務実績評価 政独委 2次評価・年度評価意見 〔事務・事業の見直しにつな がる評価や財務面の評価 等〕 (17.11.14) 国立大学法人 の年度評価	※16・17 年度の見直しの成果 ①56 法人を 42 法人に統廃合 ②非公務員化(44 法人) ③事務・事業の廃止・重点化等 政独委「勧告の方向性」 〔国立博物館など 先行独法 24 法人〕 (17.11.14) 主務大臣が見直し内容を決定、 所要の措置を実施	第3回委員長懇談会 (17.4.15) 有識者会議指摘事項 (17.10.28) 「行政改革の重要方針」 (17.12.24 閣議決定) 行政減量・効率化有識者会議 の設置(有識者会議の改組) (18.1.23)
18	日本司法支援セン ター設立 (18.4.10) 中期目標期間は平成 21 年度までの 4 年間 年金積立金管理運用 (18.4.1)	各府省独法評価委員会 の業務実績評価 政独委 2次評価・年度評価意見 〔事務・事業の見直しにつな がる評価や財務面の評価 等〕 (18.11.27) 移行独法の 見直し	政独委「18 年度の見直し方針」 (18.7.18) 政独委「勧告の方向性」 (移行独法等 23 法人) (18.11.27) 主務大臣が見直し内容を決定、 所要の措置を実施	有識者会議指摘事項 (18.5.23) 行革推進法 (平成 18 年法律第 47 号) 骨太の方針 2006 (18.7.7) 有識者会議指摘事項 (18.11.21)
19	第 4 期 住宅金融支援機構 (19.4.1) 国立大学WG、 国立病院WGを開催 (19.4～) 郵便貯金・簡易生命保険 管理機構 (19.10.1)	政独委「独立行政法人の中期目標期間終了時の見直し及び業務実 績評価に関する当面の取組方針」 (19.7.11) 各府省独法評価委員会 の業務実績評価 政独委 2次評価・年度評価意見 〔既往の政府方針等を踏まえた 評価、評価の明確性〕 (20.1.31)	政独委「勧告の方向性」 (移行独法等 35 法人) (19.12.11、12.21) 主務大臣が見直し内容を決定、 所要の措置を実施	骨太の方針 2007 (19.6.19) 「独立行政法人整理合理化計 画の策定に係る基本方針」(19. 8.10 閣議決定) 第4回委員長懇談会 (19.9.7) 有識者会議指摘事項 (19.11.27) 「独立行政法人整理合理化 計画」(19.12.24 閣議決定) 【101 独法の見直し】



年度 (平成)	主な法人の設立 評価の実施体制(政独委等)	業務実績評価 (毎年度及び中期目標期間 終了時に第三者機関が実施)	事務・事業の見直し (中期目標期間終了時に主 務大臣が実施)	備 考
23	第6期	独法分科会「平成22年度業務実績評価の具体的取組について」の策定(23.4.26) ↓ 各府省独法評価委員会の業務実績評価 ↓ 政独委 2次評価・年度評価意見(23.12.9) 評価結果の明確性、既往の政府方針等を踏まえた評価、内部統制等に係る評価等	政独委「勧告の方向性」(移行独法等9法人)(23.12.9) ↓ 主務大臣が見直し内容を決定	↓ 独立行政法人の事務・事業の見直しの基本方針のフォローアップ(23.9.15 行政刷新会議) ↓ 行政刷新会議「独立行政法人改革に関する分科会」設置(23.9.15) ↓ 独立行政法人の制度・組織の見直しについて～我が国の成長に資する政策実施機能の強化に向けて～(24.1.19 行政刷新会議 独立行政法人改革に関する分科会報告) ↓ 独立行政法人の制度及び組織の見直しの基本方針(24.1.19 行政刷新会議決定) → (24.1.20 閣議決定)
24	第7期	独法分科会「平成23年度業務実績評価の具体的取組について」の策定(24.5.21) ↓ 各府省独法評価委員会の業務実績評価 ↓ 政独委 2次評価・年度評価意見(25.1.21) 評価結果の明確性、既往の政府方針等を踏まえた評価、内部統制等に係る評価等	政独委「勧告の方向性」(移行独法等27法人)(25.1.21) ↓ 主務大臣が見直し内容を決定	独法通則法改正法案(24.5.11 閣議決定)、独法通則法整備法案(24.5.11 閣議決定) 国会提出 【独立行政法人制度を廃止し、行政法人制度を創設等】 ↓ 衆議院解散に伴い廃案 ↓ 平成 25 年度予算編成の基本方針(25.1.24 閣議決定) →制度及び組織の見直しの基本方針(24.1.20 閣議決定)の凍結 ↓ 行政改革推進本部の設置及び行政改革推進会議の開催(25.1.29) ↓ 独立行政法人改革に関する有識者懇談会の開催(25.2.28)
25	第7期	独法分科会「平成24年度業務実績評価の具体的取組について」の策定(25.5.20) ↓ 各府省独法評価委員会の業務実績評価		行政改革推進会議において、独立行政法人改革に関する有識者懇談会の中間取りまとめ(25.6.5)

※総務省HP、内閣官房HP等を基に総務省行政評価局において作成

A study on the meta-evaluation for Administrative Incorporated Agencies

Kazuhito Takenaka

Ex-Executive Staff, Administrative Evaluation Bureau,
Ministry of Internal Affairs and Communication (MIC), Japanese Government
k2.takenaka@soumu.go.jp

Abstract

From April 1st, 2001, Japanese government started the Administrative Incorporated Agency system (AIA) in the part of the reformation of the central governmental organization. This system has a notion that government must make its service in the view of the social and economic sustainability but not directly. And AIA must make its service with high-quality and -efficiency, institute autonomy and transparency. Government keeps AIA high performance with dual evaluation system; Committee on AIA Evaluation in each ministry and Committee on Policy Evaluation and AIA Evaluation in MIC (CPEA).

This paper studies the way CPEA evaluates AIA performance chronologically, especially CPEA tries to shrink the load-feeling for AIA evaluation work by the evaluatees and so on.

Keywords

Evaluation for Administrative Incorporated Agency, second evaluation, load-feeling for evaluation work

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

(投稿要領の作成公表)

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

(配布先)

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

(抜刷の配付)

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

(インターネット上の公開)

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

(著作権)

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

(事務局)

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

(以上)

『日本評価研究』投稿規定

2008.9.29改訂

2003.4.18改訂

2002.3.25改訂

2001.9.9改訂

1. 『日本評価研究』(The Japanese Journal of Evaluation Studies) は、評価に関する論文、論考、調査報告等を掲載する。
2. 『日本評価研究』は、会員間の研究成果交流の場を提供し、内外における評価研究の一層の発展に資することを主目的として発行されており、原則として会員による寄稿を掲載する。なお、依頼原稿を除き、ファーストオーサーは学会員でなければならない。また、投稿は、一時に一原稿に限るとともに、他学会誌などへ二重に投稿などのない未発表のものとする。
3. 投稿された原稿は、編集委員会の責任において審査を行ない、採否を決定する。審査にあたっては、1原稿毎に2名の査読者を選定し、査読結果を参考にする。(査読者には、投稿者名を伏せて査読を依頼する。)
4. 原稿料は支払わない。
5. 『日本評価研究』に掲載された論文等は、その全文をインターネット上の本学会のホームページに掲載する。
6. 投稿にあたっては、投稿原稿が、①研究論文、②総説、③研究ノート、④実践・調査報告、⑤その他のうち、どのカテゴリーに入るかを明記する。ただし、カテゴリーについての最終判断は、編集委員会で行なう。「研究論文」は評価の理論構築あるいは慣行の理解について重要な学問的貢献となると認められる論文、「総説」は、評価の理論あるいは慣行について概観する論文、「研究ノート」は「研究論文」作成過程での理論的あるいは経験的な研究の中間的成果物に相当する論考、「実践・調査報告」は評価事業の実践あるいは評価にかかわる調査の報告、「その他」は編集委員会が独自に企画する特集に掲載する依頼原稿等である。
7. 投稿方法
 - (1) 使用言語は日本語又は英語とする。
 - (2) 著者校正は原則として第一校までとする。
 - (3) 英文原稿については、ネイティブスピーカーによる英文チェックを済ませ、完全な英文にして投稿すること。
 - (4) ハードコピー4部(A4版)を提出する。その際、連絡先(住所、Tel、Fax、Email)と原稿の種類を明記すること。掲載可と判断された原稿については、必要なリライトを経た後に、最終原稿のハードコピー2部とDOS/Vフォーマットのフロッピーを用いたTEXTファイルを提出する。その際、オリジナル図表を添付すること。
 - (5) 刷り上がりは最大14ページとする。これを超える場合は、その経費は著者負担とする。

- (6) 日本語原稿の最大文字数は以下のとおり。①研究論文20,000字、②総説15,000字、③研究ノート15,000字、④実践・調査報告20,000字、⑤その他適宜。それぞれ和文要旨を400字程度、英文要旨を150words程度、及び和文・英文でキーワード（5つ以内）を別に添付する。印刷は1ページ、20字×43行×2段（1,720字）とする。20,000字の原稿の場合、単純計算では英文要旨1ページを加えて合計13ページとなるが、図表の量によっては、それ以上のページ数となり得るので、注意すること。
- (7) 英文ではA4版用紙に左右マージン30mmをとり、10ポイントフォントを使用し、1ページ43行のレイアウトとする（1ページ約500words）。論文冒頭に150words程度のAbstractをつける。14ページでは、7,000words相当になるが、タイトルヘッド等を考慮して、最大語数を約6,000words（図表、注、文献込み）とする。図表の量によっては、ページ数が予想以上に増える場合もあり得るので、注意すること。

8. 送付先

〒140-0002 東京都品川区東品川4-12-6 日立ソリューションズタワーB 22F
一般財団法人 国際開発センター内 日本評価学会事務局
TEL: 03-6718-5931
FAX: 03-6718-1651
E-mail: jes.info@evaluationjp.org

『日本評価研究』執筆要領

2002.9.18改訂

2002.3.25改訂

1. 本文、図表、注記、参考文献等

(1) 論文等の記載は次の順序とする。

日本語原稿の場合

第1ページ：表題、著者名、所属先、E-mail、和文要約（400字程度）、和文キーワード（5つ以内）

第2ページ以下：本文、謝辞あるいは付記、注記、参考文献

最終ページ：英文表題、英文著者名、英文所属先、E-mail、英文要約（150words程度）、英文キーワード（5つ以内）

英文原稿の場合

第1ページ：Title; the author's name; Affiliation; E-mail address; Abstract (150 words); Keywords (5 words)

第2ページ以下：The main text; acknowledgement; notes; references

(2) 本文の区分は以下のようにする。

例 1（日本語）

1.

(1)

①

(2)

(3)

例 2（英文）

1.

1.1

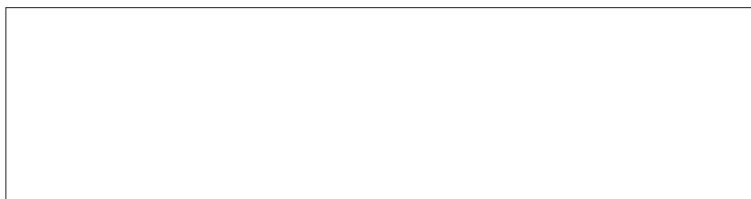
1.1.1

1.1.2

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

例 1：日本語原稿の場合

図 1 ○○州における生徒数の推移



(注)

(出所)

表 1 ○○州における事故件数

--

(注)
(出所)

例 2：英文原稿の場合

Figure 1 Number of Students in the State of ○○

--

Note:
Source:

Table 1 Number of Accidents in the State of ○○

--

Note:
Source:

(4) 本文における文献引用は、「……である (阿部 1995、p.36)。」あるいは「……である (阿部 1995)。」のようにする。英文では、(Abe 1995, p.36) あるいは(Abe 1995)とする。

(5) 本文における注記の付け方は、(…である¹。)とする。英文の場合は、(…¹)とする。

(6) 注記、参考文献は論文末に一括掲載する。

注記

1 ……。

2 ……。

(7) 参考文献は、日本語文献は著者の五十音順、外国語文献は著者のアルファベット順に記し、年代順に記載。参考文献の書き方については以下のようにする。

日本語単行本：著者（発行年）『書名』、発行所

（例）日本太郎（1999）『これからの評価手法』、日本出版社

日本語雑誌論文：著者（発行年）「題名」、『雑誌名』、巻（号）：頁－頁

（例）日本太郎（1999）「評価手法の改善に向けて」、『日本評価研究』、1（2）：3-4

日本語単行本中の論文：著者（発行年）「題名」、編者『書名』、発行所、頁－頁

（例）日本太郎（2002）「行政評価」、日本花子『評価入門』、日本出版社、16-28

複数の著者による日本語文献：著者・著者（発行年）『書名』、発行所

（例）日本太郎・日本花子（2002）『政策評価』、日本出版社

英文単行本：著者（発行年）. 書名. 発行地：発行所.

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

英語雑誌論文：著者（発行年）. 題名. 雑誌名, 巻（号）, 頁－頁.

（例）Rossi, P. H. (1999). Measuring social judgements. *American Journal of Evaluation*, 15(2), 35-57.

英語単行本中の論文：著者（発行年）. 題名. In 編者 (Eds.), 書名. 発行地：発行所, 頁－頁.

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

2名の著者による英語文献：姓, 名, and 姓, 名（発行年）. 書名. 発行地：発行所.

（例）Peters, T., and Waterman, R. (1982). *In Search of Excellence: Lessons from America's Best Run Companies*. New York: Harper & Row.

3名以上の著者による英語文献：姓, 名, 姓, 名, and 姓, 名（発行年）. 書名. 発行地：発行所.

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

（注1）同一著者名、同一発行年が複数ある場合は、(1999a)、(1999b) のようにa,b,cを付加して区別する。

（注2）2行にわたる場合は2行目移以降を全角1文字（英数3文字）おとして記述する。

『日本評価研究』査読要領

日本評価学会 『日本評価研究』編集委員会

2005年9月10日決定

1. 本査読要領の趣旨

本査読要領は、『日本評価研究』における掲載論文等の審査の要である査読手続きについて、投稿する会員及び査読を依頼される会員に対して解説を行い、審査手続きを効率的かつ効果的に行うことを目的として、定めるものです。

2. 査読の目的と投稿者の責任

査読は、投稿原稿が『日本評価研究』に掲載される論文等としてふさわしいものであるか否かについての判定を当編集委員会が行う上で必要とされるものです。

査読に伴って見いだされた疑問や不明な事項について、必要な場合は修正意見をつけて、修正を求めることがあります。査読は、その意味で、投稿原稿の改善に資するものでもあります。ただし、修正が求められた場合においても、論文等の内容に関する責任は著者が負うべきものであり、査読者の責に帰するものではありません。

査読者は2名で、編集委員会において学会会員の中から当該分野の専門家を選び依頼されますが、学会会員以外に依頼することもあります。

3. 査読の視点

査読は、以下の5つの視点によりますが、投稿原稿の種類によって、重点が違います。

- (1) テーマの重要性・有用度
- (2) 研究の独自性
- (3) 論理の構成
- (4) 実証法・方法論の妥当性
- (5) 評価理論・実践への貢献

- ・研究論文の査読については、上記の5項目全てに配慮する。
- ・研究ノートの査読については、上記5項目のうち、特に (1), (2), (3), (4) の諸項目に配慮する。
- ・実践・調査報告の査読については、上記5項目のうち、特に (1), (3), (5) の諸項目に配慮する。
- ・総説の査読については、上記5項目のうち、特に (3) と (5) の諸項目に配慮する。

4. 投稿に当たっての留意点

2. に掲げた査読の視点以外に、基本的な論文の完成度の問題があります。例えば、

- 論文等として体裁が整っているか、
- 執筆要領にしたがっているか、
- 簡潔明瞭に記述されているか、
- 実証的なデータは適切に位置づけられているか、
- 注や参考文献は本文と対応しているか、
- 専門用語の使用は適切か、
- 語句や文法的な誤りがないか、
- 誤字脱字はないか、
- 句読点に誤りはないか、
- 英文要約などの英文表現は適切か、(必ずしも和文要約の直訳である必要はなく、英文としてまとまっていること)
- 字数は規程に従っているか、

など、内容及び形式に関する留意点があります。

大学院生及び実務家の投稿において、論文としての体裁が整わないまま送付されている例があり、査読そのものに至らないものもあるので、しかるべき指導を受けた後に投稿されるよう強く勧めます。

5. 査読にあたっての判断事例

(1) 完成度において不十分であるが掲載を考慮できる場合

萌芽的な研究、発展が期待できる論文等は評価論の発展のためにできるだけ評価してください。

- 検証は十分とはいえないが、理論や定式化が学問の発展に有用である。
- 考察は十分とはいえないが、新たな理論の形成・促進に有用である。
- 文献調査は十分とはいえないが、研究の位置づけは明確である。
- 比較研究は十分とはいえないが、適用例としては意義がある。
- 考察は十分とはいえないが、社会的、または、歴史的に重要な事例の評価として意義がある。
- 考察は十分とはいえないが、特定の社会活動の評価として意義がある。
- 論文の構成や表現は適切とはいえないが、内容は評価できるものがある。
- 論理性は十分とはいえないが、実務上の有用性がある。
- 有意義な実践・調査報告である。

(2) 掲載を考慮するのが困難と判断される事例

- 問題意識や問題の設定が不明確。
- 基本的な用語の概念の理解や分析枠組が不明確または不適切。
- 論拠とするデータ等の信頼性が乏しい。
- 論旨の明確さや論証の適切さがない。
- 論文の構成、表現（用語、引用、図表等）が適切でない（または整合性がとれていない）。

6. 判定

掲載についての判定は以下の4つの類型に分かれ、最終的に常任編集委員会において決定します。ただし、これらの判定は、評価できる項目や問題のある項目の多少によるものではありません。(3)及び(4)にあるように、投稿論文の種類以外であれば、掲載を考慮できるとする場合があります。別の種類となる場合、字数の関係で、大幅に修正を要することがあります。

- (1) 掲載可とする。
- (2) 小規模の修正による掲載可とする。
- (3) 大幅な修正による掲載可とする。
但し、(総説／研究論文／研究ノート／実践・調査報告)として掲載を考慮できる。
- (4) 掲載不可とする。
但し、(総説／研究論文／研究ノート／実践・調査報告)として掲載を考慮できる。

Publication Policy of the Japanese Journal of Evaluation Studies

Last revised on 15th February 2005

The Purpose and the Name

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

Editorial Board

2. The editorial board administrates editing of evaluation study based on the editorial policy stated below.
3. The editorial board is formed with less than 20 members of the evaluation society who are assigned by the board of directors. Terms of editors are two years but can be extended.
4. The editorial board assigns one editor-in-chief, two vice-editors-in-chief, and a certain number of standing editors among the members.
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 primarily published to provide opportunities for members of the Japan Evaluation Society (hereinafter referred to as "members") to exchange findings, and to contribute to further development of the study of evaluation both domestically and internationally. As a principle, this journal publishes the contributions submitted by the members. With the exception of requested papers, the first author must be a member. A submission (as the first author) is limited to one manuscript that has not been published or submitted in any form for another journal of academic association etcetera.
3. Adoption judgments of the manuscript are made at the discretion of the editorial board. Comments from two referee readers who are appointed for every manuscript are referred to in the screening process (the editorial board requests referee readers without notifying the author of manuscript).
4. Payment for the manuscript is not provided.
5. Papers published in "The Japanese Journal of Evaluation Studies" are released on the Internet at homepage of this academic society.
6. Regarding submission, manuscripts must be identified as one of the following categories: 1) article, 2) review, 3) research note, 4) report, and 5) others. However, the final decision of the category is made by the editorial board.
 "Article" is considered as a significant academic contribution to the theoretical development of evaluation or understanding of evaluation practice.
 "Review" is a paper which provides an overview of evaluation theory or practice.
 "Research note" is a discussion equivalent to the intermediate outputs of a theoretical or empirical study in the process of producing an "article".
 "Report" is the study report related to a practical evaluation project or evaluation.
 "Others" are manuscripts for special editions requested by the editing committee.
7. Manuscript Submission
 - (1) Manuscripts may be written in either Japanese or English.
 - (2) Correction by the author is only for the first correction.
 - (3) English manuscripts should be submitted only after the English has been checked by a native speaker.
 - (4) Submit four hard copies (A4 size) of the manuscript. Contact information including mailing address, telephone number, fax number, and e-mail address, and the category of the manuscript should be clearly stated.
 For approved manuscripts, after necessary rewriting, the author needs to submit two hard copies of the final paper as well as a text file saved on a DOS/V formatted floppy disk. Original figures, charts, and maps should be provided.
 - (5) Total printed pages should not exceed 14 pages. Any cost incurred by printing more than 14 pages must be covered by the author.
 - (6) The layout for English papers should be 30 mm of margin at left and right side, 10pt for font size, 43

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

8. Mailing address

Office of Japan Evaluation Society at International Development Center of Japan
Hitachi Soft Tower B 22nd Floor, 4-12-6, Higashi-Shinagawa, Shinagawa, Tokyo,
140-0002, Japan

Phone: +81-3-6718-5931, Facsimile: +81-3-6718-1651

E-mail: jes.info@idcj.or.jp

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 photoengraved, therefore it is important that the original chart is clear. Pictures shall be treated as figures.

Figure 1 Number of Students in the State of ○○



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

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

一般財団法人国際開発センター内

TEL: 03-6718-5931 FAX: 03-6718-1651

『日本評価研究』第13巻第2号

2013年11月22日

編集・発行

特定非営利活動法人 日本評価学会

〒140-0002 東京都品川区東品川4-12-6

日立ソリューションズタワーB 22F

一般財団法人国際開発センター内

TEL: 03-6718-5931 FAX: 03-6718-1651

印

刷

株式会社 研恒社

©日本評価学会

本誌に掲載されたすべての内容は、日本評価学会の許可なく転載・複写はできません。

Japanese Journal of Evaluation Studies

Vol. 13, No. 2, November 2013

CONTENTS

Special Issue: Ten years' Experience of the Policy Evaluation System in Japanese Central Government ~Reviews and Experience~

Special Issue: Ten years' Experience of the Policy Evaluation System in Japanese Central
Government ~Reviews and Experience~ Kiyoshi Yamaya

Ten-year trajectory of the Policy Evaluation System
~Retrospective and prospective analysis of the System~ Seiichi Arai

Performance Measurement in the Evaluation Practices of Japanese Government
~Exploring the Quality of Measurement~ Tatsuya Ono

Evaluation Use in The Policy Evaluation System:
Conceptualizing "Evaluation Use" and a Tentative Analysis of Present Practice
Hiraki Tanaka

Policy Evaluation System and Accountability in Japanese Government Kazuhisa Najima

Report

A study on the meta-evaluation for Administrative Incorporated Agencies
Kazuhito Takenaka