

ISSN 1346-6151

日本評価研究

Japanese Journal of
Evaluation Studies

Vol. 19, No. 2, March 2019

ミニ特集：評価の「利用」と「影響」を考える

巻頭言「ミニ特集：評価の「利用」と「影響」を考える」 田辺 智子

業績測定の利用と影響— 公共図書館を事例とした実証分析 — 田辺 智子

評価活動は、何を、どのように、私たちにもたらしうるのか？
— 「評価の利用・影響」に関する理論研究及び実証研究を振り返る — 益田 直子

研究論文

独立行政法人制度における実績評価モデルの構築 森田 弥生

第19回全国大会開催報告

春季第16回全国大会のご案内

日本評価学会
Japan Evaluation Society

『日本評価研究』編集委員会
Editorial Board

編集委員長
Editor-in-chief

山谷 清志(同志社大学)
Kiyoshi YAMAYA

副委員長
Vice-Editor-in-chief

西野 桂子(関西学院大学)
Keiko NISHINO

常任編集委員
Standing Editors

牟田 博光(東京工業大学)
Hiromitsu MUTA

編集委員
Editors

岩渕 公二(NPO政策21)
Koji IWABUCHI

大島 巖(日本社会事業大学)
Iwao OSHIMA

岡本 義朗(内閣府)
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

田中 弥生(大学改革支援・学位授与機構)
Yayoi TANAKA

南島 和久(新潟大学)
Kazuhisa NAJIMA

西出 順郎(明治大学)
Junro NISHIDE

林 薫(文教大学)
Kaoru HAYASHI

源 由理子(明治大学)
Yuriko MINAMOTO

事務局
Office

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

日本評価研究

第19巻 第2号 2019年3月

目次

ミニ特集：評価の「利用」と「影響」を考える

田辺 智子

巻頭言「ミニ特集：評価の「利用」と「影響」を考える」…………… 1

田辺 智子

業績測定の利用と影響

— 公共図書館を事例とした実証分析 —…………… 3

益田 直子

評価活動は、何を、どのように、私たちにもたらしうるのか？

— 「評価の利用・影響」に関する理論研究及び実証研究を振り返る…………… 19

研究論文

森田 弥生

独立行政法人制度における実績評価モデルの構築…………… 35

第19回全国大会開催報告

開催の報告とお礼…………… 49

プログラム詳細…………… 51

シンポジウム報告…………… 54

共通論題セッション報告…………… 55

自由論題セッション報告…………… 66

春季第16回全国大会のご案内…………… 70

日本評価研究刊行規定…………… 71

日本評価研究投稿規定…………… 73

日本評価研究執筆要領…………… 75

日本評価研究査読要領…………… 78

Publication Policy of the Japanese Journal of Evaluation Studies	80
Information for Contributors (For English Papers)	82
Writing Manual of the Japanese Journal of Evaluation Studies (For English Papers)	84
Referee-Reading Guideline.....	86

【巻頭言】

ミニ特集：評価の「利用」と「影響」を考える

田辺 智子

国立国会図書館

評価は果たして利用されているのか。これは、評価に関わるすべての人にとって根源的な問いである。なぜなら、評価が有効に機能するためには、評価が何らかの形で利用される必要があるからである。我が国の中央省庁や自治体、公的機関で導入された評価は、どのように利用されているのだろうか。利用されていないならば、それはなぜか。どうすれば評価がもっと利用されるのだろうか。こうした問題を明らかにすることは、評価を良好に機能させる上で極めて重要である。

しかし一方で、「利用」というのは実は難しい概念でもある。最も典型的な評価の利用は、評価結果を基に施策を改廃する、評価を予算要求の根拠資料とする等の直接的な利用であろう。しかし、同時に評価には、職員の意識改革につなげる、自らの仕事に対する気付きをもたらす等の、より間接的な効果が期待される場合もあり、それらも一種の利用といえる。つまり、評価の利用は多義的な概念であり、様々な定義・解釈の余地がある。

こうした性格を持つ評価の「利用」について整理・検討する上で一つの拠り所となるのが、評価研究の分野で発展してきた「評価の利用」(evaluation use)と「評価の影響」(evaluation influence)の理論である。その出発点は、米国で1960年代から本格的に実施されるようになったプログラム評価が、評価関係者が意図したようには利用されなかったことにある。その後、「評価の利用」についての理論的・実証的研究が蓄積される中で、利用のタイプ、利用の予測因子等が明らかにされた。さらに2000年代に入ると、利用の概念を拡張し、評価が持つ様々なインパクトを可能な限り包括的に捉えようとする「評価の影響」という概念が提起された。この理論的展開の中で強調されているのは、評価の利用を狭く捉えすぎると評価の重要な機能を見落とすことになる、という点である。「評価の影響」概念に対しては一定の批判もあるものの、あえて長く使われた「利用」という用語ではなく「影響」という用語を採用している背景には、評価のもたらすインパクトの多様性と広がり反映されている。

今回のミニ特集では、この「評価の利用」と「評価の影響」の理論に焦点を当てる。益田直子氏の「評価活動は、何を、どのように、私たちにもたらしうるのか?—「評価の利用・影響」に関する理論研究及び実証研究を振り返る」では、これまで日本で十分紹介されてこなかった評価の利用・影響の既存研究を幅広く解説している。益田氏は早くからこの問題に関心を持ち、米国留学中だった2005年に、評価の利用・影響の主要論者に直接インタビュー調査した経験を持つ。続く田辺智子の「業績測定の利用と影響—公共図書館を事例とした実証分析—」は、「評価の影響」の理論を用い、公共図書館を対象とした事例分析を試みたものである。

評価の利用や影響を研究する意義の一つは、評価がもたらす直接・間接の帰結について理解を深め、評価実務の改善につなげることにある。本ミニ特集によって、多くの読者がこのテーマに目を向け、今後の議論や研究蓄積に加わってくださることを期待する。

【研究論文】

業績測定の利用と影響 — 公共図書館を事例とした実証分析 —

田辺 智子

国立国会図書館

stanabe0509@gmail.com

要 約

本稿の目的は、日本で十分紹介されてこなかった評価影響の理論を用い、評価がどのように利用され、組織や職員にいかなる影響を与えているかについて分析を行うことである。評価影響の理論では、影響の種類として、「認知・感情的アウトカム」、「モチベーション的アウトカム」、「行動的アウトカム」、「一般的影響」があると整理されている。事例としては公共図書館における業績測定型の評価を取り上げ、7館の公共図書館を対象としたインタビュー調査を基に評価影響の発現状況を分析した。分析の結果、公共図書館では、職員・組織の両面において多様な評価影響が生じていることが明らかとなった。また、評価影響の発現経路としては、評価が契機となってサービスの見直し・改善が実現する「改善ルート」、年度の目標達成に向けて意識・努力が行われる「目標達成ルート」、評価を外部への説明に利用する「対外説明ルート」の3種類が観察された。

キーワード

業績測定、評価利用、評価影響、図書館評価

1. 研究の背景と目的

わが国の公的機関では、1990年代後半以降に評価が普及し、中央府省、自治体、独立行政法人、教育機関等の様々な組織で評価が導入された。その多くは、あらかじめ指標・目標を設定して行う業績測定型の評価である。では、導入された評価は、組織内でどのように作用しているのだろうか。施策の改善やアカウントビリティ向上など、評価に期待された効果は発現しているのだろうか。

評価が有効に機能するためには、評価が何らかの形で利用され、当該組織・職員に影響をも

たらすプロセスが必要となる。もし評価が有効に機能していないとすれば、その影響プロセスのどこかに隘路があり、機能発揮に至っていないと考えられる。そうした評価の影響プロセスを明らかにすることは、評価実務を見直し、改善する取組につながりうる。

本稿では、そうした問題意識のもと、「評価影響」(evaluation influence)の理論を用いて、評価が組織・職員にもたらす影響とその発現プロセスについて分析を行う。評価研究の分野では、「評価利用」(evaluation use)についての実証研究や理論化が長く進められてきた経緯があり、評価影響の理論はその延長線上で2000年代以降に

発展したものである。そこでは、評価がもたらすアウトカムとその発現プロセスを分析するためのモデルが提示されている。日本でこれらの理論を参照している文献としては、益田（2018）、菊地（2015）、田中（2013）、吉澤（2008）があるが、国内のデータを用いた実証研究は見られない。なお、評価利用、評価影響の理論はプログラム評価型の評価を対象に発展してきたものであるが、後述するように業績測定型の評価にも適用可能と考えられる。

分析の対象としては、公共図書館を取り上げる。図書館という領域は、業績測定型の評価に古くから取り組んできた歴史を持ち、評価に用いる統計の整備も進んでいる。日本においては、図書館法によって公共図書館における評価実施が努力義務とされているため、設置自治体が行う行政評価とは別に、独自に詳細な自己評価を行う公共図書館も存在する。しかし一方で、過去の調査によれば多くの公共図書館で評価結果の具体的な活用が課題と認識されており（みずほ情報総研 2009, p.40）、評価が所期の機能を発揮しているかは検証の余地がある。本稿では、こうした背景を持ち、組織規模も比較的小さい公共図書館を分析対象とすることで、評価が組織やその内部に与える影響の一例を明らかにしたい。

2. 研究の方法

本稿の分析は、以下の二段階で行う。第一段階では、評価利用や評価影響等の文献レビューをもとに、業績測定で生じうる評価影響の類型を整理し、分析の枠組みを提示する。

第二段階では、評価を実施している公共図書館を対象にインタビュー調査を行い、モデルで示された評価影響が具体的にどのような形で観察されるかを分析する。対象としては、業績測定型の評価に十分な経験を持ち、評価が有効に機能していると考えられる館を選択する。具体的には、①どのような評価影響が観察されるか、②それら評価影響がどのようなプロセス（経路）で発現するか、の2点を明らかにする。

なお、公共図書館を対象とした評価には、図書館独自の取組として行う自己評価のほか設置自治体が行う行政評価があるが、通常、行政評価のほうが指標等も少なく、自己評価の実施館では自己評価の一部を抽出する形で行政評価を行う場合もある。このため、本稿では、より包括的な評価が行われている自己評価を分析の対象とする。また、指定管理者制度やPFI方式による運営が行われている館では評価を取り巻く状況が異なると考えられることから、本稿では基本的に直営の館を分析対象とする²。

3. 分析の枠組み

3.1 評価利用の理論

米国の公共部門では1960年代からプログラム評価が普及したが、実施された評価がしばしば意図したように利用されなかったことから、「評価利用」に関する研究と理論化が1970年代から進められた。その結果、明らかにされたものの一つが、評価利用の類型である。

初期の評価利用の研究では「利用」を非常に狭く定義しており、評価が施策の改廃につながる等の目に見える直接的な利用が想定されていた。しかし研究が蓄積する中でしだいに評価利用には様々な形態があることが認識され、今日では、評価利用の主要な類型が表1のように整理されている。評価が特定の意思決定に反映される「道具的利用」(instrumental use)、評価が関係者の認識に浸透し意思決定の情報源の一つとなる「概念的利用」(conceptual use)、評価が既存の立場の正当化や組織の合理性を示すために使われる「象徴的利用」(symbolic use)である (Herbert 2014, p.390、Mark & Henry 2004, p.36)。

1990年代に入ると、評価の「結果」のみでなく、評価実施の「プロセス」に関与することが関係者の意識改革やエンパワメントをもたらす「プロセス利用」の概念が提示された。プロセス利用は、前述の道具的利用等の類型とは異なり、利用の対象が評価結果ではなくプロセスにあることを示す概念であり、表1のいずれの類型の利用でも生じるとされる (Alkin & Taut 2003,

表1 評価利用の類型

評価利用の類型	性格
道具的利用 (Instrumental use)	施策の改廃・見直しなどの特定の意思決定に、評価が直接的に反映される。
概念的利用 (Conceptual use)	直接的な行動とは結び付けられないが、評価が関係者の認識に浸透し、意思決定の情報源の一つとなる。
象徴的利用 (Symbolic use)	既存の立場の正当化や、組織の合理性を示すために評価が用いられる。

※評価の「結果」ではなく、「プロセス」が利用対象となる場合がある。(プロセス利用)
 ※外的な圧力のために評価が利用される場合がある。(強制的利用)

(出所) Herbert, J.M. (2014) "Researching evaluation influence: A review of the literature," *Evaluation Review*. 38(5), p.390 ; Weiss, Carol H. (2005) "An alternate route to policy influence: how evaluations affect D.A.R.E." *American Journal of Evaluation*, 26(1), pp.12-30 ; Mark, M.M. and Henry, G.T. (2004) "The mechanisms and outcomes of evaluation influence," *Evaluation*. 10(1), p.36 等を基に筆者作成

表2 評価利用の類型

分析レベル	認知・感情的アウトカム	モチベーション的アウトカム	行動的アウトカム	一般的影響
個人 (Individual)	・意見・態度の強さ	・個人的目標・願望	・習得したスキルの活用 ・個人の実務変容	・熟考 ・経験則 ・スキル習得
個人間 (Interpersonal)	・担当者間の規範	・命令による規範 ・社会的報酬 ・交換	・協調的な実務変容	・正当化 ・説得 ・変化の媒介
集合 (Collective)	・アジェンダ設定 ・政策志向の学習	・構造的誘因 ・市場原理	・施策の継続・改廃 ・政策変化 ・普及	・儀式主義 ・議会公聴会 ・連合の形成 ・基準設定 ・政策検討

(出所) Mark, M. M. and Henry, G. T. (2004) "The mechanisms and outcomes of evaluation influence," *Evaluation*. 10(1), p.41 を基に筆者作成

p.10)。さらに2000年代には、エビデンスに基づく政策形成への要請から、有効性が証明された施策の採用が求められる状況が生じたことから、「強制的利用」という概念が提起された(Weiss 2005)。これについても、表1の既存の利用類型と並立するというよりは、利用をもたらす背景要因として対外圧力があるか否かによる分類となっている。

以上のような類型論に加え、評価利用の研究において理論化が進められてきたものに利用の予測因子がある。どのような条件がある場合に評価利用が多く生じるかを分析したもので、多数の実証研究が行われており、代表的なレビュー論文としてJohnson et al. (2009)、Shulha & Cousins (1997)、Cousins & Leithwood (1986)がある。予測因子については本稿の分析の対象外とするが、既存研究からは、評価の質や信頼性

等の評価自体の特性、組織環境・政治環境等の施策を取り巻く状況(コンテクスト)、評価利用者の個人的属性、評価実施過程へのステイクホルダーの関与等が主要な予測因子として挙げられている。

3.2 評価影響の理論

2000年代に入ると、それまでの評価利用の概念では、評価が施策や関係者に与える多様な影響を包括的に捉えることができないという認識のもと、「評価影響」という新しい概念が提起された。評価影響は、利用という言葉が示唆する意図的・即時的・直接的な利用に限定せず、長期的・間接的な影響を含め、評価のインパクトを可能な限り広く捉えようとする概念である。評価影響の概念を初めて提示したのはKirkhert (2000)であり、それまでの評価利用より幅広い

評価影響の性格を、「影響の源」（結果かプロセスか）、「意図」（意図的か非意図的か）、「時間」（短期か長期か）の3種類の軸でマッピングし概念整理を行った。

さらに、Henry & Mark (2003) と Mark & Henry (2004) は、評価影響を分類するのみでなく、評価が利用され何らかの変化をもたらすプロセスを表現したモデルを提示した。このモデルでは、評価がもたらすあらゆる帰結のうち、何らかの社会改善 (social betterment) につながる変化を評価影響として整理している。Mark & Henry (2004) で示された評価影響は表2のとおりである。表中、「アウトカム」は何らかの評価影響が生じたことを示すものとされており、「認知・感情的アウトカム」(cognitive and affective outcome)、「モチベーション的アウトカム」(motivational outcome)、「行動的アウトカム」(behavioral outcome) の3種類に分類されている。認知・感情的アウトカムは考え方や見解に生じる変化、モチベーション的アウトカムは目標・願望や報酬・処罰への反応にかかわる変化、行動的アウトカムは実務や政策変更など行動面に生じる変化である。「一般的影響」(general influence) とは、それ自体が直接、社会改善につながるものではないが、認知・感情面、モチベーション面、行動面に何らかの変化を起こさせることによって間接的に評価影響をもたらすものである。表2では、認知・感情的アウトカム、モチベーション的アウトカム、行動的アウトカムの順に記載されているが、アウトカムがこの順で発現するとは限らず、各種のアウトカムの間には相互作用があると想定されている。また、評価影響を分析するレベルは、「個人」、「個人間」、「集合」の3レベルに整理されている。

この評価影響のモデルと前述の評価利用の類型との関係については、認知・感情的アウトカムが概念的利用、行動的アウトカムが道具的利用、象徴的利用が一般的影響の一部（正当化と儀式主義）に対応している (Mark & Henry 2004, p.44)。

なお、このモデルは一般的な分析枠組みとして提示されており、分析対象となる評価の文脈に合わせて改定して用いることが想定されてい

る。表2で挙げられた項目は主に米国のプログラム評価の場合の評価影響を想定したものと考えられ、分析対象が異なれば異なる項目がリストアップされることになる。

3.3 その他の関連する理論・研究

(1) 業績測定の利用

以上で見た評価利用、評価影響の理論は主にプログラム評価型の評価を対象に発展してきたものであるが、業績測定で産出された情報の利用については、行政学の分野で別途、研究が蓄積されている。業績測定の普及の背景には、業績測定を実施さえすれば常に意思決定に利用され、業績測定の技術を高度化すれば意思決定への利用が促進されるという暗黙の前提があったが、それに対する疑問から実証研究が行われてきた (Taylor 2011, p.1316)。そこでは、業績情報がどの程度利用されているか、また利用の予測因子は何か分析されており、主な実証研究として、Taylor (2011)、Moynihan & Pandey (2010)、Ho (2005)、Jules & Holzer (2001) 等がある。

これらの研究では、評価利用や評価影響の理論は引用されておらず、利用の定義や類型についてはまとまった議論が行われていない。このため、業績情報の利用状況については主に行政官等の個人を対象に調査されているものの、そこで想定されているのが道具的利用なのか、「参考にする」等の概念的利用も含むのかは明らかではなく、調査対象者によって利用のイメージや含める範囲に相違が生じている可能性も否定できない。

また、利用の予測因子としては、業績測定自体の特性、政治環境、組織文化、調査対象者の個人的属性、ステイクホルダーの関与等があることが明らかになっており、論者によって整理の仕方や強調する側面は異なるものの、おおむね評価利用の研究で明らかにされてきた予測因子と共通性が高い。プログラム評価とは異なる業績測定に固有の予測因子として挙げられているものとしては、業績測定が戦略計画や目標設定に関連付けられる度合がある (Ho 2005)。

(2) 図書館評価における利用の議論

次に、図書館情報学においては、図書館を対象とした評価が「図書館評価」として研究対象となってきた(田辺 2016)。そこでも評価の利用について議論した文献は存在するが、実証研究や利用に焦点を当てた理論化はほとんど見られない。

むしろ、本稿との関係で重要となるのは、図書館評価の文献で、評価の目的やインパクトがどのように想定されているかである。図書館の業績測定についての代表的な文献では、その目的として、サービス改善、図書館の有効性についての財政当局と市民への説明、組織学習の3点が挙げられている(Poll & Te Boekhorst 2007, p.21)。一点目のサービス改善については、評価を具体的なサービスの見直しにつなげるという道具の利用を想定したものと理解できる。二点目の財政当局と市民への説明についてはアカウントリティ目的ということになるが、図書館評価の場合に特徴的なのは、単に説明責任のためというよりは、図書館の有効性を示すことで予算や支持の獲得につなげるというアピール目的が強調されていることである。これは、評価利用の類型で言えば象徴的利用と親和性が高いが、予算獲得という具体的目的のために評価が利用される場合は道具的利用と見ることもできるだろう。最後の組織学習は、スタッフがサービスや目標の達成方法についてよりよく理解し、図書館の機能についての気付きを与えることとされており(Poll & Te Boekhorst 2007, p.22)、概念的利用に当たるものといえる。

3.4 分析の枠組み

本稿では、評価影響のプロセスを包括的に表現し、評価利用の知見も織り込まれたMark & Henry (2004) のモデルを分析のベースとする。もともと評価利用や評価影響の理論は、主にプログラム評価型の評価を対象に発展してきたものであるが、以下の理由から、業績測定にも適用可能と考える。まず、評価利用の代表的なレビュー論文の一つであるCousins & Leithwood (1986) では、利用対象となる評価情報をデータ・解釈・提言等を含むあらゆる情報と定義し

ており、評価プロセスのどの時点のものでもよいとしているため(p.332)、性質的には業績測定型のモニタリングを含みうる。また、Mark & Henry (2004) の評価影響のモデルは、事例として業績測定の影響プロセスを取り上げるなど(p.48)、明示的に業績測定にも適用可能なモデルとして示されている。これまで蓄積された評価影響の実証研究を見ても、業績測定型の評価を対象としたものが含まれる³。

本稿では、表2に示したMark & Henry (2004) の評価影響モデルを、本稿の分析対象に合わせて改定した分析枠組みを用いる。表2のモデルは、評価影響が組織の範囲を超えて発現する場合にも適用可能なものとなっているが、本稿では掘り下げた分析を行うため、分析対象を公共図書館という組織の範囲内に限定する。また、表2では分析レベルが「個人」、「個人間」、「集合」の三層とされているが、本稿では分析枠組みの過度な複雑化を避けるため、「個人」に当たるレベルとして「職員」、「個人間」と「集合」を合わせたレベルとして「組織」という二層に整理を行う。

この結果、業績測定の評価影響は、「認知・感情的アウトカム」、「モチベーション的アウトカム」、「行動的アウトカム」の3種類のアウトカムと「一般的影響」が、「職員レベル」と「組織レベル」の二層について生じると整理された⁴。本章では、これらの各カテゴリについて、具体的にどのような影響が観察されるかを明らかにしていく。

4. 業績測定の評価影響

4.1 インタビュー調査

(1) 調査の概要

本稿では、業績測定を実施する公共図書館で生じている評価影響を明らかにするため、インタビュー調査を実施した。対象館は、自己評価の導入後5年以上が経過し、文献その他の情報源から評価が一定程度、有効に機能していると推測される館とした⁵。文献調査を基に4館を抽出し、スノーボール・サンプリングでさらに3館を追加

表3 調査対象館

	設置者	職員数	外部評価	調査対象者
A 館	都道府県	110人	無	5人
B 館	都道府県	70人	有	2人
C 館	都道府県	100人	無	2人
D 館	政令指定都市	200人～	無	3人
E 館	中核市	90人	有	2人
F 館	中核市	130人	有	3人
G 館	市町村	20人	無	2人

※外部評価を実施している館は、いずれも図書館協議会によるもの。

※職員数は『日本の図書館 統計と名簿 2016』（日本図書館協会、2017）による。すべての分館・地域館の専任と非常勤臨時を合計し、館の特定を避けるため概数で示した。

（出所）筆者作成

した。対象館7館の属性は表3のとおりである。

調査は2016年9～11月に実施し、半構造化インタビューによりデータを収集した。調査対象者は評価取りまとめ担当者を中心とし、可能であればその他の管理職者や一般職員を加えた。最終的な調査対象者は合計19名となった。インタビュー内容は調査対象者の同意を得てICレコーダーに記録し、内容分析のために逐語録を作成した。

なお、調査に当たっては筑波大学図書館情報メディア系の研究倫理審査の承認を受け、調査対象者に対しては研究趣旨や個人情報取扱い等について説明し同意書を取得することで、十分な倫理的配慮を行った。

(2) 対象館の評価の概要

対象館で実施されている自己評価の概要は以下のとおりである。まず、7館中5館は、評価実施の前提として3～5年程度の期間を対象とした中期計画を持ち、その実現のために年度ごとにPDCAサイクルを回していた。その他の館のうち、1館は特に中期計画を持たずに毎年度の評価を実施、もう1館は10年程度の長期の計画を持ち不定期に評価を実施していた。

全7館が評価のための指標を設定しており、毎年度評価を実施する6館は年度の目標を設定していた。評価に当たっては、既存統計に加えイベ

ント実施回数等について追加のデータ採取が行われていた。ほとんどの館で利用者アンケートを実施していたが、毎年度実施する館から隔年、不定期の館まで頻度は様々であり、アンケート結果を評価指標とする館もあれば特に評価と利用者アンケートを連動させていない館もあった。また、評価の枠組みとは別に、他の公共図書館と統計数値を比較するベンチマーク資料を作成し、年報やウェブサイト等で公表している館が3館あった。

取りまとめた評価結果については、ウェブサイトに掲載する、館内で掲示・配布する、市民グループとの懇談時の説明資料とする、主な指標の実績値についてプレス発表する等の形で公表していた。また、館長の諮問機関として設置されている図書館協議会に自己評価の結果を踏る形で外部評価を実施している館が3館あった。

ほとんどの館で、指標・目標値や評価案は係長などの実務担当者レベルで原案を作成し、ボトムアップで取りまとめを行っていた。その作業に一般の職員がどの程度参加しているかは、館により、また担当者により異なっていた。館によっては、全職員の参加のもとで計画や目標の作成が行われていた。評価制度の導入後に評価方法の見直しを行っている館が多く、指標・目標値の設定方法を変えたり、簡素化・省力化を図ったりしていた。

4.2 評価影響の発現状況

逐語録をもとに、評価影響に該当する発言を抽出し、「認知・感情的アウトカム」、「モチベーション的アウトカム」、「行動的アウトカム」、「一般的影響」の4種類のカテゴリに区分した。共通性のある発言はグループ化し、【課題の認識】、【事業見直し】等、その評価影響の性格を表すコードを付与した（文中、【 】内に記載したものがコードである）。これによって明らかになった評価影響の種類を表4に示す。表5は、各評価影響に該当する代表的な発言の例である。

以下、評価影響のカテゴリごとに分析結果を示す。文中の項番は表5の該当箇所を示したものである。

表4 公共図書館の自己評価で想定される評価影響

分析レベル	認知・感情的アウトカム	モチベーション的アウトカム	行動的アウトカム	一般的影響
職員	・課題の認識 ・業務への理解向上 ・全館的視点 ・経営的視点 ・自館の客観視	・改善意識(職員) ・目標達成意識(職員) ・対外説明意識	・目標達成努力(職員)	・対外説明力向上 ・改善策の検討(職員)
組織	・課題の共有	・改善意識(組織) ・目標達成意識(組織)	・事業見直し ・業務分析 ・計画的運営 ・目標達成努力(組織) ・対外説明への利用	・改善策の検討(組織) ・対話・意見交換 ・ライン機能

(出所) 筆者作成

表5 インタビューで観察された評価影響の例

レベル	項番	発言	コード
認知・感情的アウトカム			
職員	1	(外部評価での課題等の指摘について) やはりこれだけいろいろ書いていただくと、当然、担当者は自分達の部分を読むので。自分達、担当者でも何となく思っているが、人から言われるのはまた違う。	課題の認識
	2	(利用者アンケートによって) 利用者にもこういう希望がある、これは無理、これは誤解がある、等の気付きはあったと思う。	業務への理解向上
	3	個人としては、「図書館全体で」という考え方が少し広がったような気がする。それまでは配属された地域館のことをまず考えて仕事をしていたが、これとこれ、この館とこの館とを一緒にやったらどうか等、合理的な進め方、考え方が少し出てきたかなあと思う。	全館的視点
	4	地域館の仕事は、目の前にどんどん流れてくるものを処理するという流れにちなんで、これ(評価)があることによって、じゃあ今年何をやるのか、今年の〇〇図書館はどこに向かっていくのか、とは考えるようになった。	経営的視点
	5	客観的に自館の位置というか、そのあたりの認識に今役立っているというか。そもそも、うちがだめなところばかり見えていたが、いやいやここは全国的にもなかなかいいじゃないかという部分に気付けたというのが一つあるかもしれない。	自館の客観視
組織	6	こういう(評価の)中で気になった部分を課題として挙げていくことで、次の時にはある程度反映できないかな、みたいなところはあ。	課題の共有
モチベーション的アウトカム			
職員	7	評価は全職員には浸透していないと思う。…ただ、やはり担当者や係長クラスは意識するのではないかなと思う。私は(評価取りまとめの)担当なので一番意識するような立場にあるので、…日頃、仕事をしていても、あそこをこうするといいか、と。別に評価のためだけにじゃなくて。図書館のサービスをよくするのに、ここ工夫できそうかな、というところを意識できるようになったというか。	改善意識(職員)
	8	ちゃんと明確に目標を立てるという中では、やはり皆それぞれ気にする。その中では、「じゃあこれちょっとがんばろうか」みたいな、そんなことは考える。	目標達成意識(職員)
	9	…ともすると図書館の司書は内向きなところがあつたと思うが、そんな自分達に気付いたというのがあ。 (評価の) 作業をする中で…外部に説明するという意識が強くなったと思う。	対外説明意識
組織	10	職員は皆日々、自分がやっている仕事をよくしたいと漠然とは思っているが、それがなかなか具体化されなところがあった。そこで(評価で)目標を立てて、これをやるんですと宣言したことによって、やらざるを得なくなった。…皆が意識してそれに向かってやっていこうというのがあるかなあと思う。	改善意識(組織)
	11	(評価によって) 振り返って次に生かそうという風潮にはつながっている。	改善意識(組織)
	12	地域館ごとに目標がある。(評価で設定した目標は) 対象になる館が目標に入れて、各館の意識づけにおいて上位に設定される。	目標達成意識(組織)
	13	(係長が統計数値を確認し) 今年度のこの様子だと去年並みにならないよ、とか。…朝礼のときに報告してくれたりとか。…明確な数値目標があったりすると…最後になる前に、今このぐらい行ってるからもうちょっとがんばろうよ、みたいな話はしてたと思う。	目標達成意識(組織)
行動的アウトカム			
職員	14	…その人(個々の指標の担当者)が(年度途中で実績値を確認する際に)数字を出してきて、やばい、がんばらなきゃ、と思うきっかけになっているのだと思う。	目標達成努力(職員)
組織	15	間接的だが、私たちがサービス事業をやる中で、やり方の工夫にはつながっていると思う。関心のある利用者へ訴えかけるお知らせの仕方、何か催しをやる場合も、漫然と同じ形でやるのではなく、あちこちでやられていることと連動するとか…。	事業見直し
	16	バリアフリー…案内表示などを部門別に色分けして表示するようにしたりした。そういった気になっていた部分は反映している。	事業見直し
	17	今年度は、分析・PRチームを立ち上げて、市内の利用状況などの分析を行った。分析・PRが大事というのは中期計画の中で気付きました。新規にマンションが建った地域で登録率が低い等。そこに対する取組が必要だね、とか。	業務分析 (課題の共有)

レベル	項番	発言	コード
組織	18	確実に業務改善のツールとして機能してきているのは事実。…組織全体に、大なり小なりマネジメントのツールとしては、この仕組み(評価)は有効だったと思う。	計画的運営
	19	毎月、月ごとの統計をまとめて、その最後にこれ(指標の実績値)が載っている。…本来なら今30%以上ないといけないといった場合は、幹部会議などで、公開講座の参加者数だったら広報をもっと打てとか、HPだったら新規コンテンツを作れとか、そういうような方針が出て。…毎月、会議の後に言われたりはしている。	目標達成努力(組織) (ライン機能)
	20	半年たったら必ず副館長にその実績を示して、…少ないところはちょっと恥ずかしいなという思いを持って、実行というか、尻に火が付くとかいうのがある。	目標達成努力(組織)
	21	(職員は評価を)しっかりご覧になる。なので、評価は当然必要なことなのだなとやりながら感じる。…(評価が無かったら)今のように…この市の図書館が存続していない可能性大だと思う。委託なり指定管理なり。そういう形で押し切られてしまっていたかもしれない。…「見える化」しておくのは当然必要ということ。	対外説明への利用
一般的影響			
職員	22	(評価の)作業をする中で、外部への説明をする力が付いた…。	体外説明力向上
組織	23	外部評価で出た意見を各担当に投げで一応の答えは作ってもらっている。「それだとできない」ということもあるし、「これなら改善できる」というものもあるし。	改善策の検討(職員)
	24	レファレンスの件数とかも…。伸ばすために受付方法を検討するとか…。もっとサービスを向上させるためにどうするか、みたいなものを、この目標の数値をもとに、どこが弱いのか、と、自分達の振り返りをすごくした記憶がある。	改善策の検討(組織)
	25	(評価について職員が話題にすることはあるかという質問に対し)ありますよ、それは。…上司がハッパをかけるとか。あるいは、自発的に担当職員や班の職員が話をするとか。セクションをまたいで話すこともあるだろうし。研修などだと、研修の計画を立てる課がセクションの壁を超えて話をすると思う。	対話・意見交換
	26	…そこの班長なり課長なりは、やはり達成しないとイケないで、特定の担当にムチを入れたりとかいうことはある。	ライン機能

(出所) 筆者作成

(1) 認知・感情的アウトカム

認知・感情的アウトカムとは、評価のプロセスや結果が認識面に与える変化である。インタビューからは、評価が個々の職員レベルの認識に影響をもたらしている例が観察された。それは、いくつかの館で「気付き」という言葉で表現されていた。まず、外部評価の指摘等がきっかけで、自館の課題や業務の性格について新たな気付きが生じる【課題の認識】【業務への理解向上】が観察された(項番1,2)。そのほか、所属する分館・地域館を超えて組織全体の状況を認識できるようになる【全館的視点】(項番3)、自館の方向性を客観的な視点で見られるようになる【経営的視点】(項番4)、他館と比較した自館の相対的な状況を把握できるようになる【自館の客観視】(項番5)が観察された。

そうした気付きのうち、職員間で一定程度共有され、組織レベルの認識変化と見なすことができたものに【課題の共有】があった。具体的には、評価の過程で広報の不足や蔵書分析の必要性等の個別の課題に気付き組織的に共有されていた例(項番17)、「気になった部分を課題として挙げていくことで、次の時にはある程度反映できないかな、みたいなのところはある」(項番6)といったように評価サイクルの中で継続的に

課題を抽出・共有していた例があった。

(2) モチベーション的アウトカム

モチベーション的アウトカムとは、評価がモチベーションや意識付けに与える影響である。インタビューからは、まず、評価が契機となって職員にサービス・業務の改善が意識される【改善意識(職員)】が観察された(項番7)。また、評価によって「皆が意識して」改善に取り組むようになったという発言も見られ(項番10)、職員個人のレベルにとどまらず組織的にある程度の広がりを持って改善への指向が共有される【改善意識(組織)】も観察された。

次に、評価において年度の目標値を設定することにより、担当者がその達成を意識する【目標達成意識(職員)】が生じていた(項番8)。これについても、職員レベルにとどまらず組織レベルでも観察され、各目標の担当部署を明確にする、部門目標に評価の目標を反映させる等の形で組織的に目標達成の意識付けが生じる【目標達成意識(組織)】が見られた(項番12,13)。

さらに、評価の取りまとめ担当者には、評価の作業をする中で、外部に対し自館のことを説明しようとする【対外説明意識】が生じていた(項番9)。

(3) 行動的アウトカム

行動的アウトカムとは、評価によって生じる具体的で目に見える変化である。職員レベルでは、目標に対して年度途中の達成状況が遅れている場合に、残りの期間で実績を上げるよう努力する【目標達成努力（職員）】が観察された（項番14）。

組織レベルでは、様々な行動的アウトカムが発現していた。まず、評価に基づいてサービス・事業の改善が行われる【事業見直し】の例が多数観察された（項番15, 16）。具体的には、案内表示の見直し、広報の改善、レファレンスサービスの改善、図書館員向け研修事業の見直し等である。事業見直しの前段階として、レファレンス業務や地域ごとの登録率等を分析する【業務分析】も行われていた（項番17）。評価が継続的な改善のツールとして機能し、【計画的運営】が実現しているとした館も見られた（項番18）。また、職員レベルと共通するものとして、年度途中で進捗が遅れが見られる際に組織的に目標達成を目指した取組が行われる【目標達成努力（組織）】が観察された（項番19, 20）。

組織レベルでは、このほか、評価を外部主体への説明に利用する【対外説明への利用】が観察された。行政内部、具体的には設置自治体の図書館担当部署に対する説明に評価を用いる例が見られたほか、対市民では、ほとんどの館が評価結果をウェブサイトで公表しており、館内に掲示する、来館者が自由に持ち帰れる資料を用意する、市民団体との懇談に利用する等の様々な形で評価結果が対外的な説明資料として用いられていた。対議会では、議員が評価に関心を持っており、説明に利用するという館があった（項番21）。

【対外説明への利用】と類似するものとして予算要求に評価を利用することも考えられるが、インタビューでは評価そのものを予算要求に利用した例は見られず、中長期計画やベンチマーク資料等が利用されていた。

(4) 一般的影響

一般的影響とは、評価によって職員・組織にもたらされる変化であり、ここまで見てきたよ

うなアウトカムの発現を媒介するものである。表2のMark & Henry（2004）のモデルでは既存の社会科学や行動科学の知見から一般的影響に当たる例を抽出しているが、本稿ではこのモデルで例示された一般的影響を参考に、それに類似した性格を持つものや、他のアウトカムを媒介するという一般的影響の定義に合致するものを抽出した。

表2では、個人レベルの一般的影響の例としてスキル習得が挙げられている。インタビューからは、評価の作業にかかわることで外部に対し自館のことを説明する力が付いたとする【対外説明力向上】が見られた（項番22）。

また、表2で挙げられている政策検討（Policy consideration）に類似する一般的影響として、【改善策の検討（職員・組織）】があった（項番23, 24）。評価の過程で認識・共有された課題に対し、取り得る改善策を検討するものであり、担当者が個人的に行う例も、組織的に行う例もあった。検討するだけで終わればサービス・業務の改善は実現しないが、それが【業務分析】や【事業見直し】につながることで改善が実現することになる。

このほか、評価の様々な過程で職員間に生じる【対話・意見交換】についても、【課題の共有】や【目標達成努力（職員・組織）】等の他のアウトカムを媒介する一般的影響の性格を持っていた（項番25）。

最後に、評価で設定された目標の達成を目指す過程で、上司が部下に対し指示や働きかけを行う【ライン機能】についても、一般的影響の一つと解釈できる。部下が人事目標を設定する際に評価の目標を参照するよう上司が助言することで【目標達成意識（職員）】につながっている例があったほか、年度途中で進捗が遅れが見られる際に幹部会議で進捗を早めるよう指示が出たり（項番19）、担当者に進捗を急がせたり（項番26）することで、【目標達成努力（組織）】というアウトカムにつながる例が見られた。

4.3 評価影響の発現経路

観察されたアウトカムには、【改善意識（職員）】と【改善意識（組織）】のようにレベルを

またいで共通するアウトカムが発現しているもの、【目標達成意識（組織）】と【目標達成努力（組織）】のように異なるタイプのアウトカムが連続性や前後関係を持って発現しているものが見られた。インタビュー結果からは、こうした複数のアウトカムが関連し合っで一連の経路を形成している様子が観察された。

観察されたルートは「改善ルート」、「目標達成ルート」、「対外説明ルート」の3種類であり、それぞれの模式図を図1～3に示す。図中、実線で示した矢印はインタビューで関係性・連続性が観察されたもの、点線で示した矢印はインタビューでは明確に示されなかったものの論理的に関係性・連続性が想定できるものを表している。

(1) 改善ルート

「改善ルート」は、評価が契機となってサービスの見直し・改善が実現するルートである（図1）。

まず、評価にかかわることによって、職員レベルでは【課題の認識】、組織レベルでは【課題の共有】という類似の認知・感情面のアウトカムがレベルをまたいで発現していた（矢印①）。これら2種類のアウトカムは相互に影響しあっていると考えられる。評価書は多くの場合、ボトムアップで作成されるため、現場の担当職員が評価の過程で認識した課題は、全館的な評価書に反映されることで組織的に共有されていた（矢印②下向き）。その評価結果を見た他の職員が、担当外の業務の課題を認識することもありうると考えられる（矢印②上向き）。

次に、モチベーション面では、【改善意識（職員）】と【改善意識（組織）】という共通のアウトカムが、同様にレベルをまたいで発現していた（矢印③）。インタビューでは、「振り返って次に生かそうという風潮」（項番11）という表現が見られるなど、評価によって組織内にサービス・運営の改善を指向する雰囲気が生じており、そうした指向が職員個人に波及する可能性がある（矢印④）。

そして、こうした課題の認識・共有、改善意識は、それぞれ、サービスをよりよくするため

の行動的アウトカムにつながっていた。評価の過程で広報が課題と認識され地区ごとの登録率を分析した例（項番17）では、【課題の共有】が【業務分析】につながっていた（矢印⑤）。評価が契機となってサービス・事業の実施方法を工夫しているという例（項番15）では、【改善意識（組織）】が【事業見直し】につながっていた（矢印⑥）。そうした取組の結果、実際にサービス内容が改善したとする例も見られた（矢印⑦）。

以上のように、改善ルートでは、評価の実施に伴って認知・感情面とモチベーション面に変化が生じ、それが行動面のアウトカムにつながっていた。

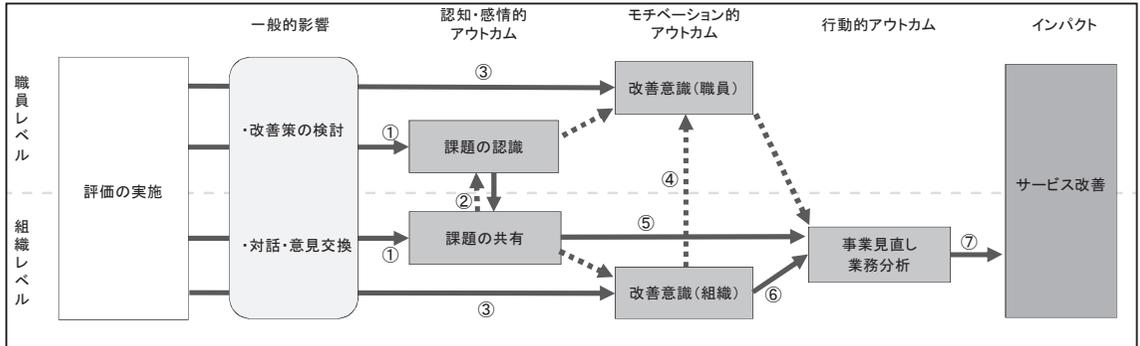
(2) 目標達成ルート

「目標達成ルート」は、評価の中で年度の目標を設定している館で観察されたルートである（図2）。明確な目標があることで、その達成が意識され、時に目標を達成するための努力が集中的に行われていた。

このルートでは、まず、モチベーション面で【目標達成意識（職員）】、【目標達成意識（組織）】という共通のアウトカムが職員・組織の両レベルで生じていた（矢印①）。人事評価の目標設定時に、係長が係員に対して評価の目標を取り入れるよう声かけするなど、組織レベルから職員レベルに働きかける作用が観察された（矢印②）。

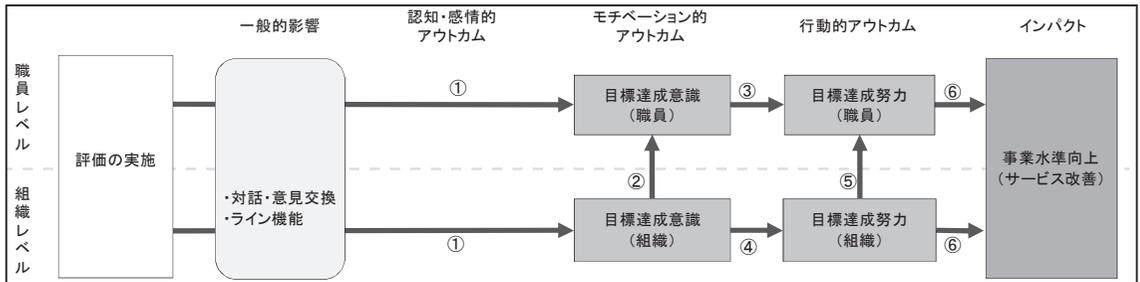
この目標達成意識は、年度を通じて常に意識されるわけではないが、年度途中で実績値を確認し、進捗に遅れが見られる場合に、特に顕著に関係者の意識にのぼっていた。現場の担当職員には、遅れがあると「やばい、がんばらなきゃ」（項番14）という意識が生じ、残りの期間での達成が目指されていた（矢印③）。組織レベルでは、幹部会議等の場で遅れが報告されると、「少ないところはちょっと恥ずかしいな」という思いを持って、実行というか、尻に火が付く…」（項番20）といったように目標達成へ向けた行動がとられていた（矢印④）。この結果、上司が部下に対し「もうちょっとがんばろうよ」（項番13）等の声かけで意識付けを与えたり（矢印②）、「特定の担当にムチを入れたり」（項番26）といったように（矢印⑤）、組織レベルから職員レベ

図1 改善ルート



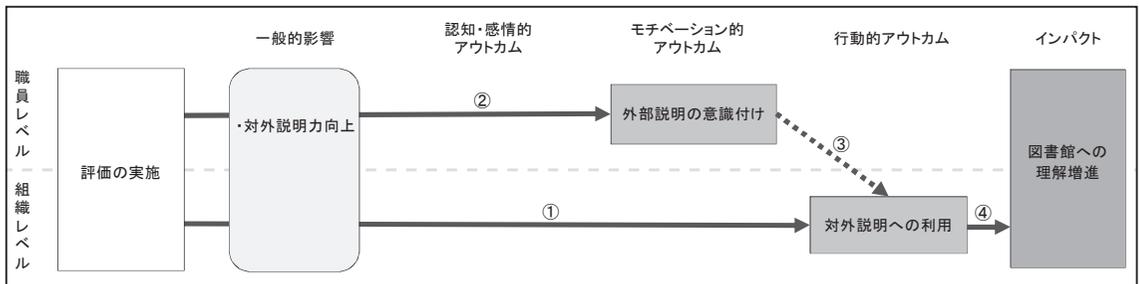
(出所) 筆者作成

図2 目標達成ルート



(出所) 筆者作成

図3 対外説明ルート



(出所) 筆者作成

ルへの働きかけも行われていた。

この結果、年度半ばでは進捗が遅れ目標達成があやぶまれていたにもかかわらず、その後の努力で目標が達成されたという例も見られた(矢印⑥)。もし評価という形で目標を設定していなかったら、低い実績値で終わっていた可能性が高く、評価があることが事業や活動の水準を高めることにつながっていた。

(3) 対外説明ルート

「対外説明ルート」は、評価を対外説明に利用し、それがアカウントビリティ向上や図書館への理解増進につながるルートである(図3)。前の2ルートと異なり、多くのアウトカムを含む複雑なルートが形成されているわけではないものの、評価による説明が実際に公共図書館に対する理解増進につながったとみられる例もあった

ため、一つのルートとして提示する。

評価は、行政内部、議会、市民への説明に様々な形で利用されていた(矢印①)。一方で、評価は、特に評価の取りまとめを担当する職員に、自館のことを外部に説明する意識付けを与えており(矢印②)、それがさらに評価を対外説明に利用する機会の増加や、有効な対外説明につながっている可能性がある(矢印③)。

このように、評価を対外説明に利用しているという時点で、評価が一定のアカウントビリティ向上につながっていると見なすことも可能であるが、それが実際に図書館に対する理解や支持に影響している例も観察された。ある館では、評価に対する議員の関心が高く、議員が評価をよく読んでおり、評価が無かったら「今のように…この市の図書館が存続していない可能性大だ」と思う。委託なり指定管理なり。そういう形で押し切られてしまっていたかもしれない(項番21)という発言が見られた。図書館への指定管理者制度の導入の是非については議論のあるところであり、本稿の検討の対象外ではあるが、少なくともこの館の場合は、職員が考える直営による存続の意義を議員に理解してもらう上で、評価が役立っていると認識されていた(矢印④)。

5. 考察

5.1 公共図書館における評価影響

(1) 評価影響の発現状況

以上の分析結果をまとめると、まず、どのような評価影響が観察されたかという点については、表4に結果が集約されている。職員・組織レベルともに、認知・感情的アウトカム、モチベーション的アウトカム、行動的アウトカム、一般的影響のすべてが観察されており、評価が職員・組織に様々なタイプの影響をもたらしていることが示された。

表4からは、認知・感情的アウトカムについては職員レベルで種類が多いのに対し組織レベルでは少なく、逆に行動的アウトカムについては組織レベルで種類が多く職員レベルで少ないことが読み取れる。認知・感情的アウトカムが組

織レベルであまり観察されなかった一因は、評価影響が一部職員にとどまり、組織全体に必ずしも波及していないことがあると考えられる。インタビューでは、「評価は全職員には浸透していないと思う。…ただ、やはり担当者や係長クラスは意識するのではないか(項番7)」という発言が見られ、評価による認識面の変化は、評価書の作成を行う担当者・係長や、その報告を受ける管理職・幹部職員を中心に発現していた。このため、【課題の認識】については一定程度、組織内で共有されているのが観察されたものの、その他の認知・感情的アウトカムについては評価に直接かかわる一部職員に限定されていたと見ることができる。一方、行動的アウトカムが職員レベルであまり観察されなかったのは、行動面で生じる目に見える変化の多くは、個々の職員が実施するというよりは、組織内の検討や意思決定を経て実施に移されるためと考えられる。

(2) 評価影響の発現プロセス

評価影響がどのようなプロセスで発現するかという点については、図1～3に示す3種類の発現ルートが明らかになった。これらのうち、改善ルートと目標達成ルートは、相互に排他的ではなく、重なり合う面がある。たとえば、改善ルートの【事業見直し】と目標達成ルートの【目標達成努力】で行われる取組が、同じものである場合もありうる。イベントの参加者数を増やすために開催日を工夫するなどのケースである。両ルートの性格を比較すると、改善ルートが目標設定の有無など評価の実施方法にかかわらず発現し、短期的にも中長期的にも機能しうるルートであるのに対し、目標達成ルートは評価で目標を設定している場合に比較的短期の年度の評価サイクルの中で機能するルートとなっている。両ルートは共にサービス改善というインパクトにつながりうるが、目標達成ルートは特に年度のサイクルの中で事業水準を上向きに押し上げることでサービス改善につながりうる性格を持っている。

観察された3種類のルートは、図書館評価の文献で想定されている評価の目的とよく合致する。

3(3)でまとめたように、図書館評価の文献では、①サービス改善、②財政当局・市民への図書館の有効性の説明、③組織学習の3点が評価の目的として挙げられている。観察されたルートのうち、改善ルートと目標達成ルートは①のサービス改善、対外説明ルートは②の財政当局・市民への説明という目的に対応している。③の組織学習については、特定のルートに対応するわけではないが、評価によって職員にもたらされる様々な気付きが組織学習につながっていると見ることができる。

以上から、図書館評価の文献で想定されている目的と合致するような評価影響が、実際にも生じているといえる。

5.2 評価利用の類型

(1) 観察された利用類型

本稿では、評価のもたらす影響を幅広く捉えるため評価影響の理論枠組みを用いて分析を行った。では、評価の「利用」という観点で見た場合、どのような利用の類型が見られるのだろうか。

評価利用の類型は、表1のように整理されている。このうち、概念的利用については認知・感情的アウトカムに包含されると考えられるが、実際、表4で認知・感情的アウトカムに分類された各種の認識変化は概念的利用と見なすことができる。これに加え、一般的影響のうち【改善策の検討（職員・組織）】についても概念的利用と見なすことができる。道具的利用と象徴的利用については行動的アウトカムに含まれると考えられ、表4の行動的アウトカムのうち【事業見直し】、【業務分析】、【計画的運営】については道具的利用、【対外説明への利用】については象徴的利用に当たると解釈できる。

これに対し、表4の残りの評価影響については、既存の評価利用の類型にはうまく当てはまらない。行動的アウトカムの【目標達成努力（職員・組織）】については、道具的利用とも象徴的利用とも性格的に異なる。またモチベーション的アウトカムに分類した評価影響については、職員の内面に生じる変化ではあるものの、概念的利用にはうまく合致しない。

つまり、本稿の分析では、道具的利用、概念的利用、象徴的利用のすべてが観察されたが、一方で既存の利用類型には合致しない評価影響も観察されたことになる。

(2) プロセス利用と強制的利用

評価利用を分類する上では、利用の対象が評価の結果なのかプロセスなのかという軸も存在する。この点でいえば、表4に示した評価影響のほとんどは、職員が評価の作業に直接かかわる中で生じたものであり、プロセス利用に分類できる。結果の利用に当たるものとしては【対外説明への利用】があるほか、評価に直接関与しない職員が評価書を見ることで気付きを得るようなケースも結果の利用に該当すると考えられる。

このようにプロセス利用が中心となっている理由は、対象館の評価が基本的に内部評価であり職員自身が評価作業を行っていること、業績測定型の評価で毎年度、評価サイクルを回すため職員が継続的に評価に関わっていることにあると考えられる。

もう一つ、評価利用の分類では、利用が外部からの強制によって生じる「強制的利用」という概念も存在する。日本の図書館法においては、自己評価を実施して改善につなげるのが努力義務とされているものの、義務付けではないため、「強制的利用」に当たる状況ではないと考えられる。また、対象館において、設置自治体から評価の利用が求められている例は見られなかった。

6. 結論

本稿では、これまで日本では適用されてこなかった評価影響の理論を用いて、公共図書館における業績測定型の評価の影響を分析した。その結果、一般に評価に期待されがちな直接的・道具的な利用にとどまらず、様々な評価影響が生じることが明らかとなった。本稿で分析対象としたのは7館という限られた公共図書館のデータであるが、評価に一定の経験を持つ館を対

象としており、またスノーボール・サンプリングで対象を増やす過程では新たな評価影響の観察が次第に減少していたことから、公共図書館で生じる評価影響をかなりの程度、捕捉できたものと考ええる。

しかし、このことと、公共図書館で評価のアウトカムが十分発現しているかどうかは別の問題である。本稿はあくまで質的分析であり、評価が有効に機能した場合に潜在的に生じるアウトカムのリストを提示したに過ぎない。全国の公共図書館で各アウトカムが実際の程度発現しているか、また相対的にどのタイプのアウトカムが発現しやすいかなど、量的な分析を行うのは今後の課題である。さらには、評価影響の予測因子を分析し、どのような変数が評価影響の発現を促進あるいは阻害するかを明らかにすることも残された課題である。評価影響を研究する意義の一つは、評価がより大きなインパクトをもたらすよう、評価実務を改善する手がかりを提供することにある。本稿が、そうした評価実務の改善に向けた実証研究の第一歩となれば幸いである。

最後に、本稿では公共図書館を分析対象としたが、他の領域の評価について評価影響を分析し、比較・検討を行うことも今後求められる作業である。自治体の一部門として、評価を予算要求等に使用しようとする公共図書館と、各施策への予算配分に評価を使用しようとする自治体の企画部門では、評価影響にも自ずと違いが生じるものと考えられる。自治体、国、公的機関等、様々な主体、コンテキストで評価影響の分析を行うことにより、評価影響にかかわる知見が一層深まることが期待される。

謝辞

調査にご協力くださった公共図書館の皆様には感謝申し上げます。

注記

1 本稿で述べた見解は筆者個人のものであり、所属機関の見解を示すものではない。

- 2 指定管理者制度の館についても比較のためにインタビュー調査を行っており、別途、分析を行う予定である。
- 3 例えばLehtonen (2012) 等。
- 4 なお、Mark (2011) では評価影響の追加のカテゴリとして「関係性の帰結」(relational consequences)が追加されており、行動や態度ではなく継続的な関係性・構造・組織プロセスに生じる変化であるとされている。本稿では組織内の限定的な範囲を分析対象とすることから、関係性の帰結については独立したカテゴリとして設定していない。
- 5 対象館の運営形態は基本的に直営であるが、一部業務を委託している館、分館の一部が指定管理者制度となっている館が含まれる。

参考文献

- 菊地端夫 (2015) 「評価情報の多義的利用に向けて：政策評価・行政評価と組織学習、市民の信頼」、『評価クォーターリー』、(32)：2-13
- 田中啓 (2013) 「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」、『日本評価研究』、13(2)：37-52
- 田辺智子 (2016) 「図書館評価と行政評価：評価方法の共通点と相違点の分析」、『図書館情報メディア研究』、14(1)：21-39
- 田辺智子 (2017) 「公共図書館における評価の“利用”」、第65回日本図書館情報学会研究大会発表、2017.11.5 (愛知県日進市)
- 益田直子 (2018) 「「評価の利用・影響」に関する理論研究の概説」、日本評価学会春季第15回全国大会発表、2018.5.26 (東京)
- みずほ情報総研 (2009) 『図書館の自己評価、外部評価、及び運営の状況に関する情報提供の実態調査報告書』
- 吉澤剛 (2008) 「評価の利用における6つのモード」、『日本評価研究』、8(1)：125-138
- Alkin, M. C. and Taut, S. M. (2003). “Unbundling evaluation use,” *Studies in Educational Evaluation*, 29, 1-12.
- Appleton-Dyer, S. et al. (2012). “Understanding evaluation influence within public sector partnership: A conceptual model,” *American Journal of Evaluation*. 33(4), 532-546.

- Cousins, J. B. and Leithwood, K. A. (1986). "Current empirical research on evaluation utilization," *Review of Educational Research*, 56(3), 331-364.
- Henry, G. T. and Mark, M. M. (2003). "Beyond use: understanding evaluation's influence on attitudes and actions," *American Journal of Evaluation*, 24(3), 2003, 293-314.
- Herbert, J. M. (2014). "Researching evaluation influence: A review of the literature," *Evaluation Review*, 38(5), 388-419.
- Ho, A. T. K. (2005). "Accounting for the value of performance measurement from the perspective of midwestern mayors," *Journal of Public Administration Research and Theory*, 16(2), 217-237.
- Johnson, K. et. al. (2009). "Research on evaluation use: A review of the empirical literature from 1986 to 2005," *American Journal of Evaluation*, 30(3), 377-410
- Julnes, P. D. L. and 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.
- Kirkhart, K. E. (2000). "Reconceptualizing evaluation use: an integrated theory of influence," *New Directions for Evaluation*, (88), 5-23.
- Lehtonen, M. (2012). "Indicators as an appraisal technology: Framework for analysing the policy influence of the UK Energy Sector Indicators," *Sustainable development, evaluation and policy-making: theory, practise and quality assurance*, 175-206.
- Mark, M. M. (2011). "Toward better research on - and thinking about - evaluation influence, especially in multisite evaluations," *New Directions for Evaluation*, (129), 107-119.
- Mark, M. M. and Henry, G. T. (2004). "The mechanisms and outcomes of evaluation influence," *Evaluation*, 10(1), 35-57.
- Moynihan, D. P. and Pandey, S. K. (2010). "The big question for performance management: why do managers use performance information?" *Journal of Public Administration Research and Theory*, 20, 849-866.
- Poll, R., and Te Boekhorst, P. (2007). *Measuring quality: performance measurement in libraries* (Vol. 127). Walter de Gruyter.
- Shulha, L. M. and Cousins, J. B. (1997). "Evaluation use: Theory, research, and practice since 1986," *Evaluation Practice*, 18, 195-209.
- Taylor, J. (2011). "Factors influencing the use of performance information for decision making in Australian state agencies," *Public Administration*, 89(4), 1316-1334.
- Weiss, C. H. (2005). "An alternate route to policy influence: how evaluations affect D.A.R.E." *American Journal of Evaluation*, 26(1), 12-30.

(2019.1.30 受理)

The Use and Influence of Performance Measurement: Empirical Analysis of Public Libraries

Satoko Tanabe

National Diet Library
stanabe0509@gmail.com

Abstract

This paper analyzes how evaluations affect people and organizations based on the theory of evaluation influence. Qualitative data were collected by interviews with librarians in seven public libraries which conduct performance measurement.

The analytical framework of evaluation influence illustrates four types of influences; cognitive and affective outcomes, motivational outcomes, behavioral outcomes, and general influences. Various influences in all types were observed in the interviews both at individual and organizational levels. Also, three types of influence pathways were observed; the improvement route, the goal attainment route, and the explanation route.

Keywords

Performance measurement, Evaluation use, Evaluation influence, Library evaluation

【研究論文】

評価活動は、何を、どのように、私たちにもたらしうるのか？ — 「評価の利用・影響」に関する理論研究及び実証研究を振り返る

益田 直子

拓殖大学

nmasuda@ner.takushoku-u.ac.jp

要 約

日本の行政府における評価活動は、制度化が進み広く取り組まれている。しかし、評価の利用及び影響に関する研究は数少ない。そもそも、「評価活動は、何を、どのように私たちにもたらしうるのか」という問いに関連する先行研究の成果が、日本ではあまり知られていない。本論文は、利用概念から影響概念への発展の説明、及びその影響概念を用いた実証研究及びレビュー研究の成果を検討することにより、先の問いへの回答を探ることをねらいとしている。「何を」（評価が与える変化）に関し利用の分類化が進展する一方で利用が自己目的化するという課題を克服する為、影響概念は、評価の目的の射程を「社会の改善」にまで延ばした。また、従来あまり取り組まれてこなかった「どのように」（変化をもたらす要因と変化の間の連鎖）を包括的に把握しようとする取組が、各国で始まっている。今後、日本の経験が影響の経路の探求に寄与していく事を期待したい。

キーワード

評価の利用、評価の影響

1. はじめに¹

日本において、評価活動は日常的な行政活動の一部になったと言っていいだろう。中央省庁においては政策評価制度が導入され、地方公共団体においてはその約6割²において行政評価が導入されている。加えて、2018年現在、首相官邸においては、行政改革推進本部の下で進める行政事業レビュー、まち・ひと・しごと創生本部の下で進めるKPI（重要業績評価指標）、そしてIT総合戦略本部の下で進めるEBPM（証拠に基づく政策立案）の取組があり、また、財務省においては予算執行調査の取組がある。事業や施策

を評価し、その結果を活用するための仕組みは増えた。このように、異なる要請に応答するために政府における活動領域を拡大してきた評価活動は、何を、どのように、私たちにもたらしうるのだろうか。

評価の利用に関する研究を牽引してきたキャロル・ワイス（Carol H. Weiss）は、1960年代の変革の時代に、アメリカで社会政策が拡大するに伴い活発化する評価活動に初めて携わった当時を振り返り、次のようにインタビューに答えている³。最初の評価結果に対する反応はどのようなものでしたか、という質問に対して、彼女は、「ほとんどなにもなかった」と答えている。

また、その評価結果は3年をかけて三分冊にまとめられ評価依頼者である政府に40部送付されていたが、「だれもその箱を開封していなかった」、そして、「反応の欠如が、その後の私の研究の多くの拠り所となった」と答えている。こうした問題意識が広く共有されるようになり、1970年代以降、評価研究において「評価の利用」が熱心に研究されるようになった。そして2000年代に入り、新たに「評価の影響」という概念がそれを継続し発展させるものとして現れた。

本論文では、評価研究における、評価の利用から影響概念への発展と、影響概念に基づく分析モデルの説明及び実証研究への適用について先行研究の内容をレビューすることにより、評価活動は、何を、どのように、もたらしうると考えられてきたのかを明らかにすることを目的とする。諸外国における研究の進展が日本に比べ先行している現状から、諸外国における研究の紹介が中心となる。これらの紹介により、日本における評価活動の取組の広がり、何を、どのように、評価実施者・組織・社会等にもたらしているのかを考え、議論するきっかけの一つになることを期待している。

本論文の構成は次の通りである。大きくは理論研究と実証研究に分ける。前者の理論研究部分については、さらに評価の「利用」と「影響」に分ける。またそれぞれを、各概念の説明と、その概念に関連する研究のレビュー論文の説明に分ける。他方、後者の実証研究部分では、「評価の影響」理論を用いた研究事例の紹介を行う。そして最後に、それらの先行研究から何を読み取り、議論しうるかについて論ずる。

2. 理論研究—「評価の利用」(evaluation utilization / use⁴⁾)

(1) 「評価の利用」の概念

「評価の利用」について理解を深めようとする動きは、1970年代から顕著になった。1960年代後半の米国において社会政策が拡大し、様々な方面から、実際に想定した効果をもたらしたのか、という疑問が投げかけられた。そこで政府

は本格的に評価を行うようになったが、実際には評価書の多くは政策過程ではほとんど使われなかったという背景がある(益田 2010)。

この時期に定義づけられ、評価の利用の分類として広く知られているのは、道具的(instrumental)、概念的(conceptual)、象徴的(symbolic)利用⁵の3つ⁶である。道具的利用とは、直接的、具体的であり、それらは観察可能な利用である(Rich 1977)。決定の支援及び課題解決の機能を持つ。例えば、評価結果に基づき事業への予算配分を決定する場合を挙げることができる。

概念的利用とは、評価結果は共通の知識に取り入れられ、意思決定者の思考枠組みの一部を形成するようになる、という利用である。それは必ずしも直接的に明白な行動変化をもたらす必要はない(Rich 1977, Weiss & Bucuvalas 1977)。言い換えれば、教育的機能を持つ。例えば、評価結果が、意思決定者が持つ事業についての考え方を修正した場合を挙げることができる。

そして、象徴的利用とは、隠された動機や自己の利益のための利用であり(Knorr 1977, Leviton and Hughes 1981, Greene 1988)、政治的機能を持つ。象徴的利用の他に、正当化のための利用、政治的利用、説得のための利用等、さまざまな呼ばれ方をしている。例えば、立法への支持を得るために評価結果を利用する場合を挙げることができる。

その後も新たな分類が提起された。例えば、1990年代に入り提起された、プロセス利用(process use)を挙げることができる。評価への参加経験は、評価書又はその勧告と同じく、行動に影響を与えうるという考えを表している(Patton 1997)。上記3分類は評価結果の利用(use of evaluation findings)に関する分類である一方、プロセス利用は、評価に参加する過程によって生ずる変化に焦点を当てた分類である点が異なる(Alkin & Taut 2003, Weiss et al. 2005)。2000年代に入り、「評価の利用」は「評価の影響」に発展していくが、その発展に寄与したKirkhart (2000)は、プロセス利用と他の利用の分類とを統合した3次元モデルを提示した。だが、「評価の影響」は、多くの点でワイス(Weiss 1988)の

評価に関する見解を反映している、と考えられている (Herbert 2014)。この詳細は、本論文の「3. (1)『評価の影響』の概念」で説明する。

また、影響概念への発展につながる他の研究として、次のレビュー論文の項目で紹介をする Cousins & Leithwood (1986) がある。同論文は、「利用」に関する65の研究の体系的調査を行い、決定等に直接的に利用されたかに主に着目する「利用」概念に対し、直接的な利用よりもっと捉えがたい効果が存在することを明示したことから、のちの影響概念の創出につながる「文脈 (context)」の重要性が高まることとなった。

(2)「評価の利用」に関するレビュー論文

「評価の利用」に関する議論は、1970年代以降増加していった論文を体系的にレビューし、膨大な研究成果を統合する作業 (メタ分析等) を繰り返しながら展開してきた側面がある。その代表的な取組が次の4つの論文、Leviton & Hudghes (1981)、Cousins & Leithwood (1986)、Shulha & Cousins (1997)、そして Johnson et al. (2009) である。

各論文の概要については分量が多いため、ここでは、本論文の研究課題である「評価活動は、何を (従属変数としての「評価の利用」)、どのように (評価の利用に至るまでの各変数間の相互関係)、私たちにもたらしうるのか」を明らかにすることを意識しながら、各論文が明らかにしたこと、及び明らかにしていないことについて、筆者自身の考察を行う。

Leviton & Hudghes (1981) では、利用をもたらした要因を、関連性、コミュニケーション、情報の加工、信頼性、利用者の関与と支持の5つに分類しているが、それらの各要因がどのような利用をもたらしたのかについては局所的な扱いである。言い換えれば、「利用」という従属変数と、その要因と考えられる独立変数との関係は十分に検討されておらず包括的には分からない。各要因が5つの分類の中で利用に対して肯定的に作用、否定的に作用、又は場合によって両方の作用があることを明らかにし列挙するに止まる。つまり、従属変数と独立変数の関係について、独立変数の収集と列挙は行われているものの、

従属変数 (「何を」) が十分に特定化されていない。また、評価の利用に至るまでの各変数間の相互関係 (「どのように」) に関する包括的検討もない。

次の Cousins & Leithwood (1986) では、Leviton & Hudghes (1981) と比べ、従属変数をより明確に定義し、独立変数との関係を特定化することを試みている。さらに、文脈や背景からの影響を考慮した評価の利用の検討を行っている点が特徴的であり、その後の評価の利用に関する研究において文脈や背景の議論を活発化させた (Shulha & Cousins 1997)。しかし、独立変数と従属変数の関係は基本的に単線的であり、同論文が自ら課題として指摘しているように、各変数間の相互関係が十分に検討されていない。つまり、「何を」の特定化は進展したけれども、「どのように」までは十分に検討できていない。

次の Shulha & Cousins (1997) は、上述の2論文とは異なり、独立変数と従属変数の収集と統合を試みしていない。評価者とプログラムの利害関係者がともに評価情報を生み出し、普及し、利用することへの責任を共有するという関係変化が生じている現状認識を持ち、新たな理論的・実証的研究が求められているという考えから、評価の利用に関する研究をレビューしている。その結果、プロセス利用という新たな用語の必要性、評価者の役割の多様化、文脈と評価活動及び評価の利用の密接な関係性、短期ではなく長期にわたる評価結果の追跡による包括的な把握の必要性、誤用についての対話の促進の必要性等、評価研究の新たな方向性を提示している。しかし、「どのように」に関する理論又は実証的モデルの検討までは行っていない。

そして、Johnson et al. (2009) では、Cousins & Leithwood (1986) の分析枠組みを用い、1986年以降に公表された評価の利用に関する実証論文のレビューを行い、再び、独立変数と従属変数の収集と統合を試みている。加えて、Shulha & Cousins (1997) の前提となっていた評価活動における変化に配慮し、新たな特徴や要因を加えている。そして、レビューの結果、利害関係者の関与が、評価の利用をもたらす評価過程や環境の創出を促進することについて最も多くのレ

ビュー論文から確認できたことから、同論文は、評価利用者と評価者との間における積極的なコミュニケーションこそが、長期的にみれば評価の利用を最大化する、と強調した。しかし、「評価の利用」に関する実証研究の分析枠組みは、伝統的な3分類に見られるように、基本的には、評価結果の利用に焦点を当てている。そうしたことから、評価過程における変化（プロセス利用）、又は評価の影響過程の追跡（評価の影響）といった新たな概念的枠組みを、分析枠組みに組み込むことができていない。そのため、同論文は、自ら発見した知見を従来の「評価の利用」の分析枠組みに接ぎ木する形で示すに止まる。つまり、従来の独立変数の集積に止まり、プロセス利用や「評価の影響」といった新たな考え方を組み込むモデルの提示まではできていない。

以上の4つのレビュー論文の存在は、評価の利用に関する研究が数多く取り組まれてきたからこそ、発見された多くの知見を統合しようとする試みが度々行われてきたことを表している。そして「評価の利用」の研究と実践を発展させるために必要な事項に関する検討が積み重ねられ、その成果が次の「評価の影響」モデルに引き継がれている。新たなモデルの土台に、上記レビュー論文が対象とした膨大な研究からの知見があることを覚えておく必要がある。それは、新たなモデルが文脈、ニュアンス、包括的な因果関係の理解を特徴としていることの原因になっていると考えることができるからである。

Johnson et al. (2009) が指摘するように、2005年までのレビュー対象論文の中には、「評価の影響」モデルが示す影響の経路 (pathways) ⁷ を使って論ずることも可能に見えた論文もあったが、実際には、利用をもたらした変数の収集を行うに止まっていた。彼女らは、その理由として、影響の経路を明らかにする研究にはより創造的な取組が求められることを挙げている (p.388)。しかしその後、影響理論を用いた実証研究が発表されていった。次に、同理論の説明と、実証研究の例について紹介する。

3. 理論研究一「評価の影響」(evaluation influence)

(1) 「評価の影響」の概念

上述の「評価の利用」の箇所でも触れたように、「評価の影響」は、多くの点でWeiss (1988) の評価に関する見解を反映している⁸。特定された意思決定者に評価が役立つことが強調されてきたことに対して、Weiss (1988) は、Cronbach et al. (1980) 及びCronbach (1982) の考えを引用しながら、評価の利用を理解するために、実際の意思決定は多様な主体による多様な活動の結果である、という見方の重要性を強調した。特に大規模な政府プログラムにおいては、意思決定者は評価結果を受けたからといって、それをすべて必ず実践に移すことができるような「善意の専制君主 (the benevolent despot)」ではない、と論じた。こうした意思決定についての見方を前提として、彼女は、「評価の役割は、交渉等の話し合いを支援することが重要であり、正しい決定を指摘することではない」等の意思決定における評価の役割を示した。また、これらの見方を前提に、評価結果が影響力を得るには、それが進行中の政策的議論に受容される必要があり、その受容経路は多様であるとした。そして、こうした多様な経路を経由し、評価結果は意思決定者による政策に対する考え方に影響を与えうると論じた。このように、Weiss (1988) の「利用」の考え方は、評価は、評価が実施された場所からは離れた場所で間接的に決定に影響を与える、というものであった。この考え方は、2000年代の「評価の影響」の議論に引き継がれ、洗練化されていくことになる。

「影響」概念に関する代表的な論文として、Kirkhart (2000)、Henry & Mark (2003)、Mark & Henry (2004) の3つを挙げることができる。それらの論文では、「利用」は、直接的・意図的・即時的であり、それらはいつも起こりうるわけではなく、評価がもたらしうるインパクトを過小評価しがちであると指摘した。一方、「影響」は、間接的・非意図的・長期的な要素を含む、評価が与えうるインパクトの全体を分析しようとするものであり、また、「利用」を引き

起こしうるメカニズムや過程を理解することに着目する概念であると主張した。次に、各論文の概要について紹介する。

(2) カークハート (Kirkhart) の理論

「評価の影響」研究の嚆矢となったKirkhart (2000) は、評価の影響の範囲について、影響の源 (source)、意図 (intention)、そして時間 (time) の3つの側面から再考することを目的としている (p.5)。

Kirkhart (2000) がまず初めに説明している歴史的背景の概略は、次の通りである (pp.5-7)。評価者は自らの仕事をもたらすインパクトの特徴や範囲について、長年関心を示してきた。こうして影響に関する議論は、内部・外部評価、評価者の役割、評価の専門性、倫理と価値、結果の利用といったテーマの下で生じた。しかし、専門性が高まるにつれて、影響に関する議論がばらばらになるとともに、影響に関する理解は不安定なものになっていった。かつて利用 (utilization/use) は、データに基づく、評価結果の影響 (the data-based influence of evaluation findings) に焦点を当てていた。だが、この定義は狭いため、プロセス利用といった概念が現れた。しかしこれも、部分的解決でしかなかった。例えば、「利用」は、利用対象が評価結果ではない場合、意図しない効果が発生する場合、そして時間が経つに伴い徐々にインパクトが現れる場合には、不十分で不適切な用語であった。また、影響の履歴が、「結果に基づく利用」 (results-based use) の観点から追跡された場合、プロセス利用等の他の形の影響の履歴が捕捉されない、等の課題が残った。このように、「結果に基づく利用」という観点をベースに、そこに新たな知見を追加して「利用」の概念の範囲を拡大させるという方法だけでは、評価の影響を明確に把握することはできない。その克服には、影響について多角的な地点から検証するという方法で再概念化すること、また、「結果に基づく利用」に関するこれまでの理解をより広い文脈の中で検証する枠組みをつくるが必要になる。

上述の問題意識から、次にKirkhart (2000) は、用語を「利用」から「影響 (influence)」に意図

的に移行する重要性について論じている。利用を表すuse及びutilizationは共に、単一方向的、挿話的、意図的、道具的な影響を表している。他方、影響とは、無形又は間接的な方法によって他に効果をもたらす、人々又は物事的能力又は権力である。この概念は「利用」より広く、多角的、漸变的、非意図的、非道具的である。また、この概念には「利用」が表す、単一方向的、挿話的、意図的、道具的な影響も含まれる。

そして、次にKirkhart (2000) は、意図 (intention)、源 (source)、時間 (time) の3次元から成る影響に関する統合的理論を提示している。意図 (intention) は、意図されていた (intended) 影響か、又は意図されていなかった (unintended) 影響か、の2つに区分される。源 (source) は、影響の源又は起点が、評価の結果 (results) か、又は過程 (process) か、の2つに区分される。そして、時間 (time) は、評価の影響が現れるタイミングによって、評価中 (immediate)、評価の終了時点 (end-of-cycle)、又は未来 (long-term) に区分される。

Kirkhart (2000) は、各次元について詳細な検討を行った上で、最後に、影響に関する統合的理論の効用について、9つの項目を挙げている (pp.18-20)。ここでは項目を列挙するに止める。(i) 利用に関する議論の明確化、(ii) 個々の評価を取り巻く影響の特徴を明らかにするための枠組みの創出、(iii) 時間の経過に伴い発展する影響に関するパターンの追跡、(iv) 利用 (use) と誤用 (misuse) の分類化、(v) 影響に関する研究の妥当性の向上⁹、(vi) 影響に関する研究のメタ分析の促進、(vii) 評価理論の発展の追跡、(viii) 評価理論間の比較の促進、(ix) 理論構築の支援。

そして最後に、評価の影響に関するより包括的な視点が評価の専門職集団全体にとって肯定的な影響をもたらすことを認識することが重要である、と指摘する。また、評価の影響に関する統合的理論によって、これまで私たちが理解してきた以上に、評価活動が大きな波及効果を持つことを私たちは認識するようになる、と強調する。

(3) マークとヘンリー (Mark&Henry) の理論

上述の通り、利用をめぐる課題の克服策としてKirkhart (2000) のような影響への移行を主張する議論が現れた。しかし、Kirkhart (2000) に基本的には賛同しつつも、それも不十分であり、影響についてより詳細かつ明確な分析枠組みと用語の発展が必要であると主張した¹⁰のが、Henry&Mark (2003) 及びMark&Henry (2004) である。それらは、包括的理論の構築を目指し、評価の影響に関する概略的理論を提示した (Mark&Henry 2004, p.39)。ここでは、両論文における「影響」概念に関する理論の詳細を説明していく。

①影響メカニズムの類型

まず初めにHenry&Mark (2003) では、分析レベルを、個人 (individual)、個人間 (interpersonal)、そして集合的 (collective) の3つに分け、評価が影響をもたらす変化の過程 (process) と結果 (outcome) を区別した (pp.297-305)。個人レベルは、評価の過程又は結果が、個人の思考又は行動に変化を直接に引き起こしている場合を指す。個人間レベルは、個人間の相互作用にもたらす変化を指す。そして集合的レベルは、公的又は民間の組織が行う決定と活動に、評価が直接又は間接に与える影響を指す。

表1の横軸は3つの分析レベルを示し、各分析レベルに該当する、変化の過程又は結果に関する特定の型 (specific types of change processes or outcomes) が列挙されている。例えば、表1では個人レベルの型は、Elaboration、Heuristics、Priming…と並んでいる。各分析レベルに並ぶ各型は、心理学、政治学、組織行動論等の他の研究領域からも採用している。これは、Henry&Mark (2003) が、従来の評価の利用に関する研究及び理論の発展が妨げられてきた要因の一つに、他の研究領域における、変化の過程に関する研究成果を活用してこなかったことがある、と考えていることを反映している (p.299)。

これら3つの分析レベルの理解に際しての留意点の1つ目は、ある分析レベルのある変化の過程又は結果に関する特定の型が、同じ分析レベルの別の型を引き起こす場合もあれば、別の分析

レベルの別の型を引き起こす場合もあるという点である。例えば、個人の考え方に起きた変化は個人レベルの行動の変化を引き起こすこともあれば、集合的レベルの変化を引き起こす場合もあるということである (p.298)。2つ目は、評価の影響を研究する際に、変化の過程又は結果に関するすべての型を常に考慮することを期待してはいない点である。むしろ、調査者による初期のデザインとしての利用可能ナリストとして、また、実務家が考慮に入れるものとしての利用を想定している (p.299)。

分析レベルを3つに分けてそれぞれに変化の過程又は結果に関する特定の型を明示する考え方の大事な点は、異なる結果 (outcome) がつながり、連鎖し、経路 (pathway) をつくることによって、評価の影響についての私たちの議論を次の段階に発展させることにある (p.299)。

次に、Mark & Henry (2004) では、上記のHenry&Mark (2003) が提示した3つの分析レベルという分析軸に、もう一つの分析軸として「影響のメカニズムの類型 (Type of process/outcome)」を加えた (p.40)。その結果が表1である。Henry&Mark (2003) には「影響のメカニズムの類型」の分析軸についての言及はない。

「影響のメカニズムの類型」は、一般的影響過程 (general influence processes)、認知的・感情的過程 (cognitive and affective processes)、動機づけの過程 (motivational processes)、行動的過程 (behavioral processes)、関連的帰結 (relational consequences) の5つが示されている。なお、最初の4類型はMark & Henry (2004) で示され、最後の関連的帰結はMark (2006) で加えられた類型である。

認知的・感情的過程とは、思考や感情に起こる変化を指す。行動的過程とは、行動における変化を指す。一方、一般的影響過程と動機づけの過程は、評価が促す可能性のある最終的な変化 (the ultimate change) ではなく、その変化に至るまでの経路における、初期又は中間の段階に注意を払う。一般的影響過程は、影響の経路において他の変化が生ずることを後押ししうような、比較的基本的な過程であり (Mark 2006, p.132)、他の中間的・長期的結果をもたらすかど

うかによってその重要性が判断される過程である (Mark & Henry 2004, p.41)。同様に、動機づけの過程は、賞罰への人間の反応に対する目的や

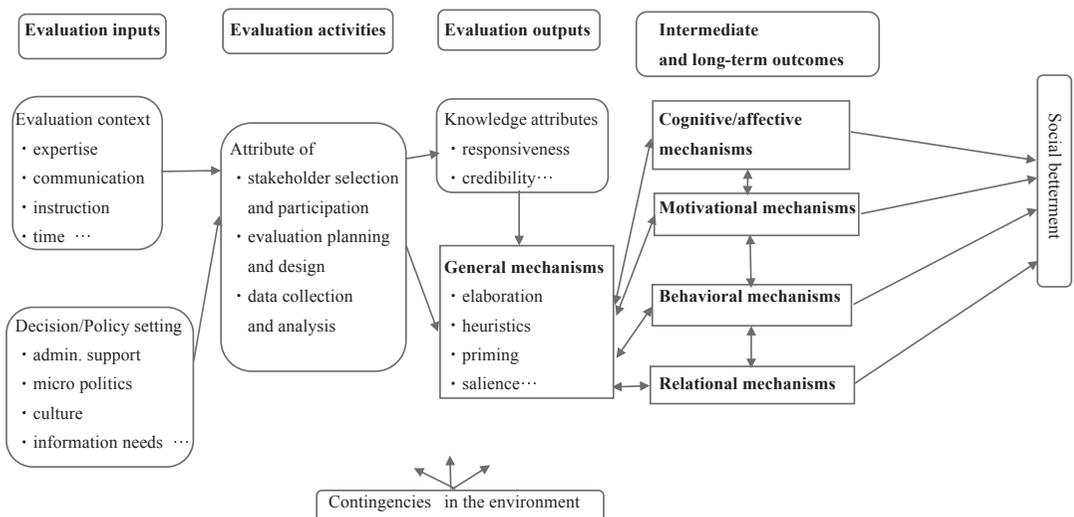
願望を指すが、これは行動に変化をもたらす上で重要な役割を果たす過程である。そして、関連的帰結は、評価者が行動又は考え方ではなく、

表1 影響メカニズムの種類

Type of Process/Outcome	Level of Analysis		
	Individual	Interpersonal	Collective
General influence	Elaboration	Justification	Ritualism
	Heuristics	Persuasion	Legislative hearings
	Priming	Change agent	Coalition formation
	Skill acquisition	Minority-opinion influence	Drafting legislation
			Standard setting
		Policy consideration	
Cognitive and affective	Saliency Opinion/attitude valence	Local descriptive norms	Agenda setting Policy-oriented learning
Motivational	Personal goals and aspirations	Injunctive norms	Structural incentives
		Social reward exchange	Market forces
Behavioral	New skill performance Individual change in practice	Collaborative change in practice	Program continuation change or cessation
			Policy change Diffusion
Relational	Self-perception of empowerment	Networks	Democratic forum
		Shifts in power relations	Learning organizations
			Social justice

(出所) Mark & Henry(2004)、Mark(2006)、Gildemyn(2014)

図1 評価の影響に関する概略的理論



(出所) Mark & Henry (2004)、Mark (2006)、Mark&Henry (2013)、Gildemyn (2014) を参考に筆者作成

進行中の関係、構造、及び組織の過程を変更しようとする取組を指す (Mark & Henry 2013, p.151)。

影響過程は、表1の各類型の異なる分析レベルを横断する形で追跡できる。例えば、行動的過程の類型を個人レベル (活動における個人的変化) から個人間レベル (活動における共同的变化) へ、そして個人間レベルから集合的レベル (政策変化) へと対応関係を追跡できる (Mark & Henry 2004, p.43)。また、表1は出発点であり、今後の理論的及び実証的研究によって修正されることが望ましい。

②評価の影響に関する概略的理論 (Schematic theory of evaluation influence)

「評価の影響に関する概略的理論」は、図1の通りである。評価の究極の目的を、「利用」ではなく、「社会の改善 (social betterment)」とし、それに向かう一連の過程を各段階の連鎖関係 (評価インプット→評価活動→評価アウトプット・一般的メカニズム→中長期のアウトカム→社会の改善) として示している。社会の改善 (social betterment) とは、社会問題の減少と人間にとって必要なものの充足 (Mark et al. 2000, p.24) を意味する。言い換えれば、民主的社会の文脈において社会の改善とは、以前あった状態よりも良いとみなされる状態をもたらすことであり、その判断は、討議と世論を通じて行われる (Henry&Mark 2003, p.295)。このように、「利用」した後にもたらされるであろう、この「社会の改善」を、研究者や評価者等に意識させる理論となっている。

この理論はCousins (2003) から着想を得ている (Mark & Henry 2004, p.37)。また、上記の表1で示した影響メカニズムの類型は、この概略的理論の重要な部分となっている。影響メカニズムの類型のうち、一般的影響過程以外の4類型は、図1の中長期的結果 (intermediate and long-term outcomes) に位置づけられる。この概略的理論は、プログラム理論又は変化の理論を評価活動に適用している。この適用は、従来の利用の議論が利用の前兆となりうる要因の特定と、利用の分類を行ってはいたものの、評価が考え方・動機・

行動に影響を与える変化の過程はほとんど考慮されてこなかったという不足 (gap) を埋めることを目的としている。その結果、今後の研究の道標になるような様々な仮説を生み出すことが可能になるだろう、と論じている (Mark & Henry 2004, p.45)。

Mark&Henry (2004) は、例として、米国の州レベルの業績測定を紹介している (pp.48-50)。この例においては、予算担当官は行政府と立法府に分かれる。立法府の予算担当官は予算の最終的決定である立法過程に関与することから、これら職員のみを対象に予算決定に対する業績測定の影響についてインタビューを行うと、「あまり影響していない」、という結果になってしまう。しかし、実際には、それ以上の影響が表れていることを、複雑な評価の影響の経路を追跡すると発見できる。具体的には次のようになる。

「予算担当官にとって業績測定結果は重要 (salient) になる [個人レベル]。→業績測定結果は行政機関の予算策定過程においてより重要になる [個人間レベル]。→当該行政機関において、より上位の中央予算査定部局に対して予算要求をする際に、その要求を正当化 (justify) するために業績測定結果を利用する [個人間レベル]。→中央予算査定部局における歳出予算の審議において、業績測定情報に基づく論拠を説得力をもって (persuasively) 利用する [個人間レベル]。→州の予算責任者が議会に提出する歳出予算案 (draft appropriations bill) の中に、業績測定の影響が具体化される [集合的レベル]。→議会の予算担当官の視点からは、法案に至るまでの業績測定による直接的影響の全てを見ることは、それらの影響が法案に組み込まれてしまっているために、できない [集合的レベル]。」よって、議会 (立法府) の予算担当官にインタビューを行っても、業績測定が「予算決定にあまり影響していない」という回答になってしまう。しかし、影響の経路を追跡すると、行政府において影響が表れていることを発見できる¹⁾。

この例を、表1の影響メカニズムの類型を用いて解釈すると、個人レベル (salient) は「認知的・感情的過程」に、個人間レベル (justification & persuasion) は「一般的影響過程」に、そして

集合的レベル(draft appropriations bill & legislation)も「一般的影響過程」に該当していることが分かる。

この理論について重要なことは、最終的な状態(例えば政策変更)に至る場合と至らない場合の両方の条件を特定することである(Henry & Mark 2003, p.306, Mark & Henry 2004, p.52)。また、メルヴィン・マークへのインタビュー¹²によれば、図1が示す一連の過程は、変更を必要とした文脈や背景を説明できるならば、分析対象に合わせて変更を加えていくことは望ましいという。

③「影響」と「利用」の関係

Mark & Henry (2004, pp.43-44)は、利用の類型と影響メカニズムの類型との関係を次のように説明している。道具的利用は行動的過程に該当する。但し、従来の道具的利用では、個人の活動における変化に至る経路と、政策変更等の集合的变化に至る経路は異なるにもかかわらず、考慮されてこなかった。一方、影響概念においてはその相違を考慮に入れている。また、概念的利用は認知的・感情的過程に該当する。但し、従来の概念的利用では、啓発(enlightenment)が起こっているのは個人、集合、又はそれらの両方においてなのかについて曖昧であった。しかし、影響概念では明確に区別している。次に、象徴的利用は、個人間レベルの「正当化(justification)」及び集合的レベルの「儀式主義(ritualism)」に該当する。そして、プロセス利用については該当する類型がない。プロセス利用は、影響が評価結果(evaluation outputs)ではなく評価活動(evaluation activities)によってもたらされたことで定義されるものだからである。

④小括

「利用」の研究において欠如又は不十分であった事項、具体的には、定義や分析レベルの不明瞭さ、最終的結果への過度な注目(Mark & Henry 2003, p.311)、それに至るまでの過程の不明瞭さといった事項を克服するために、影響理論はつくられた。具体的には、分析レベルや影響メカニズムを明確化した。また、プログラム理論、

及び他の研究領域における変化の過程に関する研究成果を活用し、分析レベルや影響メカニズムの類型を関連づけた。このようにして、影響理論は、評価が社会の改善に至るまでの包括的な影響の過程を一步ずつ(step by step)¹³又はドミノを順に倒すように¹⁴、把握することをねらいとしている。

また、上記の①と②の説明にある通り、「影響」の分析枠組みは柔軟に作られているため、多様な事例を分析対象とすることを可能にし、実際、多様な国々における多様な分析レベル及び政策領域を対象に、事例研究の蓄積と理論の改善との往復作業が進められている。

(4)「評価の影響」に関するレビュー論文

「評価の影響」に関する論文を体系的かつ包括的にレビューをしている論文は、現段階では、Herbert (2014)のみである。紙幅の関係上、ここでは同レビュー論文は何を明らかにしていないかに関する筆者の考察について論ずる。

Herbert (2014)は、評価の影響概念を用いた実証研究のレビューを行い、影響の定義、研究方法、影響概念の有効性等を検討している。その結果、定義の明確性、方法論の厳密性、調査開始のタイミングへの配慮¹⁵、そして分析結果に生じうるバイアス¹⁶等への留意を研究者に促している。

しかし、それらの課題への対応策として、評価の影響に関する概略的理論をどのように改良していけばよいのかについては論じていない。その代わりに、新たに現れた影響概念を、評価が持つインパクトを理解するための発展途上のアプローチであるとして肯定的な考えを示すとともに、研究の良事例を明らかにし、今後の研究が発展することへの期待を表している。よって、それらの事例を次に説明していく。

4. 実証研究－「評価の影響」理論を用いた研究事例

(1) 諸外国

①Herbert (2014)において選ばれた良事例

影響メカニズムの説明基準と過程を明確にしているのみならず、実際その追跡にも他と比べて成功していると考えられているのは、Oliver (2008) 及びGildemyn (2014) の2つの論文である (p.411)。まず初めにOliver (2008) は、人的援助の評価報告書が援助機関の政策と実践に与える影響を検証することを、論文のねらいとしている。分析枠組みは、Henry & Mark (2003) の「影響の経路」を使用している。また、解釈においては、学習する組織 (learning organization) の理論を踏まえ、組織が学習しないという点に注目した分析も行っている。その理由は多岐にわたるが、学習する文化の欠如、アカウントビリティの欠如といった要素を挙げている。それを受けて、評価は学習の構造 (the learning structure) の中で道具として十分に理解されていない可能性がある、と指摘している。分析手法は、①評価を行った、又は評価結果を受け取った／利用した個人へのインタビュー、②評価報告書の内容分析であり、3つの評価書の比較分析を行っている。

次に、Gildemyn (2014) の概要は次の通りである。ガーナ政府の政策のモニタリング及び評価 (M&E) は、独立した市民社会組織 (CSOs) が担っており、そのM&Eの取組効果を明らかにするのが、論文のねらいである。分析枠組みは、Henry & Mark (2003) とMark & Henry (2004) が示した理論を用い、次の2つのリサーチクエスションを立てている。1つ目は、選定された地区において、CSOが国民健康保険制度のM&Eを行うことによって、どのようなタイプの影響のメカニズムが起こったか、である。2つ目は、影響のメカニズムを引き起こすものとしての政策対話の役割とは何か、である。分析手法は、比較事例研究デザインを採用している。ガーナの北部地域から4つの地区を選定し、そのうちの2地区では政策対話の機会が設けられている一方、他の2地区では設けられていない。各会議と、鍵となる情報提供者及びCSOの職員へのインタビューによる質的分析が行われた。その結果、分析対象の地区においては、認知的・感情的過程よりも動機づけの過程の方が、行動的過程を引き起こしやすいことが分かった。具体的には、直接

顔を合わせて行われる会議で評価結果が議論される場合には、利害関係者から、問題を解決するための約束や積極的関与等を引き出すことができていた。言い換えれば、行動の透明性が確保されることによって他者から見られているという意識が促され、それが社会的賞罰として働き、正しい行動を誘発する傾向が見られることが分かった (但し、決定的な証拠に裏付けられている訳ではない)。Gildemynはこの会議における約束を「行動に向けた約束 (pledges for action)」と呼び、動機づけの過程のもとに新たな型として加えている。だが、行動に変化が生じたとしても一度きりの場合が度々観察されたことから、それを「一度だけの対応 (one-time action)」と呼び、行動的過程のもとに新たな型として加えている (pp.516-519)。一方、対話の場が設定されていることによって、新たな約束の取り付けではなく、時には、既に解決の約束がなされているにも関わらず未解決のままとなっていた課題の優先順位を上げ、解決に向かわせる圧力を強めることができることを、明らかにしている。

②Herbert (2014) によるレビュー以降の研究

Herbert (2014) がレビュー対象論文を選定した際に用いた評価の影響に関する理論的研究である、Kirkhart (2000)、Henry&Mark (2003)、Mark&Henry (2004) の3つの論文に、実証研究の良事例として紹介されたOliver (2008) 及びGildemyn (2014) の2つの論文を加え、合計5つの論文のそれぞれを、2014年1月から2017年8月末までの間に引用した論文を検索した。検索は、SAGE Journal検索によるCrossRefを基本とし、そこで検索できない論文については、Proquest又はGoogle検索による被引用論文も併せて参考にした。総計47の論文が検索結果として現れた。それらの各論文の要約を読み、「評価の影響」が研究の中心的概念に位置づけられており、かつ複数の検索結果に現れた論文を選定した。その結果、4つの論文、Alkin&King (2017)、Sridharan, et al. (2017)、Mark (2017)、Bundi (2016) が選定された。各論文の概要を、以下に紹介する。

Alkin & King (2017) は、利用 (use) と誤用

(misuse) の考え方の説明とともに、「影響」を「利用」と関連づけながら論ずることをねらいとしている。3つの論文をシリーズ化し、「評価の利用」に関する概念の発展を歴史的に説明する方法を採っている。本論文はそのうちの2番目の論文¹⁷であり、「評価の影響」概念の登場とそれに対する疑問を呈する内容で終わっている。同論文では、影響概念は、評価に関する効果を幅広く考えることを評価研究者に促した点では貢献しているが、「利用」から「影響」の用語に完全に移行することに対しては必要ではない、という立場を示している。しかし、この主張に関する詳細な議論は、今後の3番目の論文で行うという。

Sridharan, et al. (2017) 及びMark (2017) は、評価研究学術誌であるNew Directions for Evaluationの154号「健康格差を評価するための能力構築－中国・インド・チリにおける評価実験からの教訓」に収録された各論文である。不平等に与える「評価の影響」についてSridharan, et al. (2017) が行った調査に対し、Mark (2017) が影響概念からコメントをする、という構成になっているが、その大半は影響概念を説明する内容となっている。

Bundi (2016) は、スイス議会が評価を必要とする理由を明らかにすることをねらいとしている。プリンシパル・エージェント理論等を踏まえて仮説を立てた上で、アンケート調査を連邦及び地方レベルの議員に対して行った。その結果、行政府に責任を課す (accountable) ことを求める議員が評価を望むことが明らかになった。

Bundi (2016) は、影響に関する研究の制約（証拠に基づく組織として受け止められることに関心を持つ関係者による自己申告に依拠等）に対し、議会の視点からの調査は新しい洞察を与えると主張している。

以上の通り、Herbert (2014) 以降の「評価の影響」に関する研究は、それに対する「利用」概念の有効性を主張する論文、事例研究に「影響」概念の適用を試みる論文、又は従来の実証研究には希薄であった、評価実施主体ではなく、評価の利用者である議会の視点から調査することの意義を主張するために「影響」概念を用い

る論文等が現れていることが分かる。これらのことから、今後は、「影響」概念を用いた実証研究が増えより良い理論を構築しようとする動きの一方で、「利用」概念の有効性の主張が再び活性化し、両者間での議論が活発になるかもしれない。しかし両者は、評価が社会等に対して与える影響を明らかにしたいという目的は共有している。実際、「影響」概念の論者は、「影響」は「利用」を含む概念である、と主張している。また、「影響」を用いた論文が世界の様々な評価活動を対象に発表されてきており、今後もその傾向は継続すると考えられる。

(2) 日本

「評価の利用」及び「評価の影響」をキーワードとしCiNii Articles及びCiNii Dissertationsにより2017年8月末現在までの論文検索を行った。「評価の利用」での検索結果として15論文が現れた。そのうち「評価の利用」の概念が研究の中心的概念に位置づけられているのは、吉澤 (2008) 及び田中 (2013) の2論文であった。なお、「評価の影響」での検索結果は、該当なしであった¹⁸。吉澤 (2008) は、「評価の利用」についてそれまでの議論を詳しく解説し理論的整理を行い、新たに6つの分類を提示している。但し、「影響」とは区別するとともに意図的な効果に概念を限定して「利用」を分類している。一方、田中 (2013) は、日本の政策評価制度における「利用」の実態分析のためのモデルを構築する際に、Kirkhart (2000) が提示した「影響」モデルを参照している。

その他、「評価情報の利用」をキーワードとすると、佐藤 (2011) 及び菊地 (2015) も該当する。佐藤 (2011) は予算編成における財政担当部門による評価情報の利用度を決定する要因について論じている。一方、菊地 (2015) は評価情報の非道具的利用の考え方に注目し、組織学習及び市民の信頼確保のための利用について論じている。但し、これら2論文は、本論文で扱う「利用」に関する理論について簡潔に言及するに止めるか、又は明示的には言及していない。よって、管見によれば、日本においてはこれまで、「利用」についての理論的又は実証的研究につい

ては数が少ないながらも見られるが、「影響」についての研究はほとんどない。

そうした中で、本稿は、評価の利用のみならず影響についてのこれまでの理論的及び実証的研究の説明とともに、それらのレビュー論文の説明を包括的に行う初めての取組であると言って良いだろう。さらに、日本を事例とした影響理論を用いた実証研究が、学会報告という形式で現れ始めている。田辺智子は、2017年11月5日に図書館情報学会第65回研究大会において「公共図書館における評価の“利用”」という題目で発表している。また、益田は、2017年11月11日に米国評価学会（AEA）において「評価の影響－日本の立法府と行政府の比較分析（The Influence of Evaluation: A Comparison of the Legislative and Executive Branches in Japan）」という題目で発表している。今後、他国と同様に日本においても、本論文冒頭で言及したような評価活動の広がりに伴い、影響理論を用いた研究への取組が増えていくことを期待したい。

5. おわりに－先行研究に何を読み取り、議論しうるか

本論文では、評価活動は、「何を」、「どのように」、もたらしうると考えられてきたのかという2つの視点を明示し、これまでの研究成果の包括的解釈を試みた。本論文で筆者が明らかにしたことは、まず、「何を」（従属変数。利用及び影響といった、評価活動が与える変化）については、利用の分類化が発展してきたということである。しかし、新たな分類が追加されていくものの、「利用」が最終的な目的であるかのように見える段階に至った。また、利用をもたらす要因（独立変数）の収集も熱心に行われてきたが、それらと従属変数との関係は基本的には単線的で、また、その間に介在する要因（結果・過程）間の連鎖については十分に注目されてこなかった。言い換えれば、評価活動が「どのように」変化を与えるに至るのかを明らかにしてこなかった。そうした中、2000年代に入り、評価活動が社会の改善に至るまでに及ぼす作用を、3つの

分析レベルに分けて追跡し経路を明らかにすることによって、包括的に把握しようとする「評価の影響」と呼ばれる考え方が現れた。「何を」の射程が「利用」のみならず、「社会の改善」にまで伸びた。さらに、「どのように」を明らかにするために、評価研究のみならずそれ以外の社会科学領域の研究成果を積極的に活用し概略的理論が発案された。そして、まだ数は少ないものの、同理論を用いた実証研究が多様な国々を事例に始まっている。このように現段階では、「どのように」については発展初期段階にあり、ある一定の結論を導くことはできない。今後、「評価の影響」に関する概略的理論を用いた実証研究が質・量ともに向上し、研究成果の統合が可能な蓄積量になった時に、「評価の利用」と同様にレビュー研究（メタ分析等）が繰り返され、研究成果が統合され、徐々に明らかになっていくものと考ええる。

一方、「評価の影響」に関する概略的理論を用いた実証研究が進むにつれて、次の事項が課題として、今以上に広く認識され議論が求められてくるものと、筆者は考える。1つ目は、最終的な従属変数を「社会の改善」としたことによって、その概念を観察可能な変数に操作化する際にその妥当性の確保には、従来の「利用」の研究以上に注意を向ける必要が出てくるものと思われる。言い換えれば、「社会の改善」とは何かを明確に定義づける作業は、研究対象が置かれる文脈を明確かつ詳細に説明する作業を伴うことになる。2つ目は、評価研究以外で発展してきた変化の過程に関する研究成果も積極的に評価研究に取り入れる傾向が今後も続いた場合、それに伴い、各研究領域で発展してきた用語や考え方について評価研究領域内でコンセンサスを得るために、研究者や実務者間での頻繁な議論が必要になってくると思われる。

本論文の着想を得たのは、2005年開催の米国評価学会とカナダ評価学会の合同研究大会の会場及びその後のEメールで、当時米国留学中であった筆者が、評価の利用及び影響に関する代表的論者（例えば、Alkin, Christie, Greene, Henry, King, Kirkhart等）に半構造化インタビュー調査を行った時であった。当時は、米国評価学会の

「評価の利用」に関する分科会主催の研究会が契機となり、本論文でも紹介をしたKirkhart (2000) 論文を始めとする論文が『評価の利用の範囲を広げる (The expanding scope of evaluation use)』という題目でNew Directions for Evaluationに特集され、それに続いてHenry&Mark (2003)、Mark&Henry (2004) と新たな理論が次々と発表された頃であった。そのためか、インタビュー調査の結果があまりに多様で集約が難しく、新たな議論が始まったばかりという印象を受けた。本論文の執筆中、鮮明に脳裏に浮かんだのは、評価は様々な知識のうちの一つであるという認識を持ちつつも、そのうちの一つの重要なものとして社会に与える作用の実態を解明しようと互いに議論し格闘している、評価領域の研究者と実務家の姿だった。より良い社会の実現に評価は役立つはずであるとその価値を認める人々の層の厚みが、評価の影響理論の発現の根底にある文脈となっているようにも思われる¹⁹。

最後に、メルヴィン・マーク及びゲーリー・ヘンリー自らが論文で認めている通り、影響の理論的モデルは、影響を及ぼす経路が多様であるがゆえに複雑であり (Mark&Henry, 2013)、それを本論文では、可能な限り個別に深入りしすぎることなくシンプルに、しかし重要な点を網羅できるように紹介することを心掛けた²⁰が、予想通り長くなってしまった。最後までお読みくださった皆様に、感謝を申し上げる。

注記

- 1 匿名の査読者より有益なコメントを頂戴した。記して感謝の意を表したい。なお、本研究は日本学術振興会科学研究費若手研究 (B) (課題番号26780094) の助成を受けたものである。
- 2 総務省『地方公共団体における行政評価の取組状況等に関する調査結果』は、2016年10月1日現在、全地方公共団体の61.4%が行政評価を導入していると発表している。都道府県及び施行時特例市以上の市に限れば、9割以上の地方公共団体が導入している。
- 3 The Oral History Project Team (2006, p.476) .
- 4 評価研究者はuseとutilizationを同義のものとして使用している (Johnson, et al. 2009, p.378)。

- 5 Leviton & Hughes (1981) が分類の議論をしており、「象徴的 (symbolic) 利用」ではなく、同様の意味で「説得的 (persuasive) 利用」を使っている (p.528)。
- 6 4つ目の分類として、imposed use (課された利用) と呼ばれる、プログラムの利害関係者は評価結果に注意を払わざるを得ない、という利用の種類がある (Weiss, et al. 2005, p.12,16-17)。
- 7 「影響の経路 (pathway of influence)」は、評価とその最終目標である社会の改善 (social betterment) の間にある中間的なアウトカムの関係 (個人・個人間・集会的) を指している。Henry & Mark (2003) においては、変化のメカニズム (change mechanisms)、変化の過程 (change processes)、作業仮説 (working hypotheses) といった表現を用いて “pathway” を説明している。
- 8 Weiss自身は、2000年代に「影響」という用語を使う論文の現れは、Weiss (1979) の提案を受けたものであると指摘している (Weiss, et al. 2005, p.14)。
- 9 具体的には、従属変数を十分に特定すること、影響の範囲を明確に定めること等を挙げるができる。
- 10 Kirkhart (2000) との関係についての説明は、Mark & Henry (2004, p.39)、及びHenry&Mark (2003, p.309) を参照。なお、KirkhartとHenry&Markの理論間の詳細な比較分析は管見の限りない。但し筆者は、前者が米国評価学会での一連の報告を一つの理論 (Figure 1.2.) にまとめ上げる役割を持つ一方、後者はより実証研究を進めるための理論構築を目指す内容になっていると考える。
- 11 Melkers&Willoughby (2001) .
- 12 2017年11月11日に米国評価学会研究大会会場 (ワシントンDC) にて実施。
- 13 Henry & Mark (2003, p.311) .
- 14 Mark (2017, pp.134-135) .
- 15 評価開始直後の調査開始が良いという考えに対し、評価後に時間間隔を置けば、個人の変化と集会的変化の間の連関を観察しやすくなるという考えがある。影響メカニズムを最もよく把握するための、調査開始のタイミングに配慮が必要である。
- 16 研究対象の組織の関係者による自己申告 (例：インタビュー調査、組織文書の調査) に依拠する調査では、分析結果に生じうるバイアス等への留意が必要

である。

- 17 最初の論文は、Alkin, M.C. & King, J.A. (2016) . The historical development of evaluation use. *American Journal of Evaluation*, 37, 568-579.
- 18 益田 (2010) は「社会の改善」理論を用いてGAOを研究対象とする意義を論ずるに止まる (pp.77-79)。
- 19 評価文化については、益田 (2016) を参照。
- 20 Henry&Mark (2003) 及びMark&Henry (2004) は、理論枠組みの構成の根拠となる研究成果と事例を、Mark (2011) は、影響分析を行う際のヒントを、そしてMark&Henry (2013) は、理論枠組みが依拠する考えを明らかにしている。それらを読めば、彼らの理論枠組みをより深く理解するとともに、それがおお発展途上にあることも知ることができる。

参考文献

- 菊地端夫 (2015) 「評価情報の多義的利用に向けて：政策評価・行政評価と組織学習、市民の信頼」、『評価クォーターリー』、32 (1) : 2-13
- 佐藤徹 (2011) 「評価と予算の連動メカニズムの実証分析 - 予算編成過程における行政評価情報の利用度の規定要因 - 」、『公共政策研究』、11 : 71-84
- 田中啓 (2013) 「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 (特集 政策評価制度10年の経験：レビューと展望)」、『日本評価研究』、13 (2) : 37-52
- 益田直子 (2010) 『アメリカ行政活動検査院 (The U.S. Government Accountability Office: GAO) - 統治機構における評価機能の誕生』、木鐸社
- 益田直子 (2016) 「評価政策と評価文化の相互作用」『評価クォーターリー』、38 : 24-44
- 吉澤剛 (2008) 「評価の利用における6つのモード」、『日本評価研究』、8 (1) : 125-138
- Alkin, M. C., and King, J. A. (2017). Definitions of evaluation use and misuse, evaluation influence, and factors affecting use. *American Journal of Evaluation*, 38 (3) , 434-450.
- Alkin, M. C., and Taut, S. M. (2003). Unbundling evaluation use. *Studies in Educational Evaluation*, 29, 1-12.
- Bundi, P. (2016). What do we know about the demand for evaluation? Insights from the parliamentary arena. *American Journal of Evaluation*, 37 (4), 522-541.
- Cousins, J.B. (2003). Utilization effects of participatory evaluation. In T. Kellaghan and D.L.Stufflebeam (Eds.) , *International handbook of educational evaluation*, Boston: Kluwer Academic Publishers, 245-266.
- Cousins, J. B., and Leithwood, K.A. (1986). Current empirical research on evaluation utilization, *Review of Educational Research*, 56, 331-364.
- Cronbach, L. (1982). *Designing evaluations of educational and social programs*. San Francisco, CA: Jossey-Bass.
- Cronbach, L.J. , Ambron, S. R. , Dornbusch, S. M. , Hess, R. D. , Hornik, R. C. , Phillips, D. C. , Walker, D. F. , and Weiner, S. S. (1980). *Toward reform of program evaluation*. San Francisco, CA: Jossey-Bass.
- Diaz-Puente, J. M., Montero, A. C., and Carmenado, I. R. (2009). Empowering communities through evaluation: Some lessons from rural Spain. *Community Development Journal*, 44, 53-67.
- Diaz-Puente, J. M., Yague, J. L., and Afonso, A. (2008). Building evaluation capacity in Spain: A case study of rural development and empowerment in the European Union. *Evaluation Review*, 32, 478-506.
- Fjellstrom, M. (2007). The influence of evaluation in higher education: Curriculum development through a deliberative responsive dialogue with stakeholders. *Evaluation Journal of Australasia*, 7, 25-30.
- Gildemyn, M. (2014). Understanding the influence of independent civil society monitoring and evaluation at the district level: A case study of Ghana. *American Journal of Evaluation*, 35 (4), 507-524.
- Greene, J. C. (1988). Communication of results and utilization in participatory program evaluation. *Evaluation and Program Planning*, 11 (4), 341-351.
- Henry, G. T., and Mark, M.M. (2003). Beyond use: Understanding evaluation's influence on attitudes and actions. *American Journal of Evaluation*, 24, 2003, 293-314.
- Herbert, J. L. (2014). Researching evaluation influence: A review of the literature. *American Journal of Evaluation*, 38 (5), 388-419.
- Johnson, K., Greenesid, L.O., Toal, S. A., King, J. A., Lawrenz, F., and Volkov, B. (2009). Research on evaluation use: A review of the empirical literature from

- 1986 to 2005. *American Journal of Evaluation*, 30 (3), 377-410.
- Kirkhart, K. E. (2000). Reconceptualising evaluation use: An integrated theory of influence. *New Directions for Evaluation*, 88, 5-23.
- Knorr, K. D. (1977). Policymakers' use of social science knowledge: symbolic or mstrumental? In C. H. Weiss (Ed.) *Using Social Science Research in Public Policy Making*. Lexington, MA: Lexington Books.
- Lehtonen, M. (2010). *Indicators as an appraisal technology: Framework for analysing the policy influence of the UK energy sector indicators*. Retrieved from <https://www.researchgate.net/publication/286128271/download>
- Leviton, L.C., and Hudghes, E.F.X. (1981). Research on the utilization of evaluations: A review and synthesis. *Evaluation Review*, 5 (4), 525-548.
- Mark, M.M. (2006). *The consequences of evaluation: Theory, research, and practice*. Presidential address Evaluation 2006: Annual meeting of the American Evaluation Association, Portland, OR.
- Mark, M.M. (2011). Toward better research on -and thinking about- evaluation influence, especially in multisite evaluations. *New Direction for Evaluation*, 129, 107-119.
- Mark, M. M., (2017). Inequities and evaluation influence. *New Directions for Evaluation*, 154, 127-138.
- Mark, M. M., and Henry, G. T. (2004). The mechanisms and outcomes of evaluation influence. *Evaluation*, 10, 35-57.
- Mark, M. M., and Henry, G. T. (2013). Multiple routes. Evaluation, assisted sensemaking and pathways to betterment, In M. C. Alkin (Ed.), *Evaluation roots: A wider perspective of theorists' views and influences*, 2nd ed. Los Angeles, CA : Sage, 144-156.
- Mark, M. M., Henry, G. T., & Julnes, G. (2000). *Evaluation: An integrated framework for understanding, guiding, and improving policies and programs*. San Francisco, CA: Jossey-Bass.
- Melkers, J. E. and Willoughby, K. G. (2001). Budgeters' views of state performance budgeting systems: Distinctions across branches. *Public Administration Review*, 61 (1), 54-64.
- Oliver, M. L. (2008). *Evaluation of emergency response: Humanitarian aid agencies and evaluation influence* (Doctor of Philosophy in Public Policy). Georgia State University, Atlanta, GA.
- Patton, M. Q. (1997). *Utilization-focused evaluation: The new-century text*, 3rd ed. Thousand Oaks, CA: Sage.
- Rich, R. F. (1977). Uses of social science information by federal bureaucrats: knowledge for action versus knowledge for understanding. In C. H. Weiss (Ed.), *Using social research in public policy making*. Lexington, MA: Lexington Books.
- Shulha, L. M., and Cousins, J. B. (1997). Evaluation use: Theory, research, and practice since 1986. *Evaluation Practice*, 18(3), 195-208.
- Smith, N. L., & Chircop, S. (1989). The Weiss-Patton debate: Illumination of the fundamental concerns. *Evaluation Practice*, 18(3), 5-13.
- Sridharan,S., Zhao, K., and Nakaima, A. (2017). Editors' notes. *New Directions for Evaluation*, 154, 9-16.
- The Oral History Project Team. (2006). The oral history of evaluation, part 4: The professional evolution of Carol H. Weiss. *American Journal of Evaluation*, 27, 475-484.
- Weiss, C. H. (1979). The many meanings of research utilization. *Public Administration Review*, 39, 426-431.
- Weiss, C. H. (1988). Evaluation for decisions: Is there anybody out there? Does anybody care?. *Evaluation Practice*, 9, 5-19.
- Weiss, C. H., and Bucuvalas, M. J. (1977). The challenge of social research to decision making. In C. H. Weiss (Ed.). *Using Social Science Research in Public Policy Making*. Lexington, MA: Lexington Books.
- Weiss, C. H., Murphy-Graham, E., and Birkeland, S. (2005). An alternative route to policy influence: How evaluations affect D.A.R.E. *American Journal of Evaluation*, 26, 12-30.
- Willoughby, K. G., and Melkers, J. E. (2001). Assessing the Impact of Performance Budgeting: A Survey of American States. *Government Finance Review*, 17(2), 25-30.

Evaluation Use and Influence: A Review of Theoretical and Empirical Research

Naoko Masuda

Takushoku University
nmasuda@ner.takushoku-u.ac.jp

Abstract

Internationally, Japan's evaluation system is considered well-established. Research on evaluation use and influence, however, has not yet been fully developed in Japan. To fill this gap, this article reviews related studies, focusing on independent variables (causes of changes) and dependent variables (changes: use or influence). The concept of evaluation influence is able to not only expand a range of dependent variables, such as social betterment, as opposed to the use itself, as the ultimate goal of evaluation, but also focus on gaining a comprehensive understanding of relations between each independent and dependent variables using program theory. It seems difficult for social betterment to be operationalized clearly in empirical studies, but various contexts underlying evaluations in a variety of countries can be accepted. Future Japanese research will contribute to build a comprehensive theory of evaluation influence.

Keywords

evaluation utilization, evaluation use, evaluation influence

【研究論文】

独立行政法人制度における実績評価モデルの構築

森田 弥生

筑波大学大学院

morita@gssm.otsuka.tsukuba.ac.jp

要 約

平成26年の独立行政法人通則法の改正により、新たな独立行政法人制度が登場した。同制度においては「政策実施機能を最大化」するとの観点から、適切な業務の実績評価と適切な目標策定のために、政府の統一的な指針が示されている。しかし、現実には、目標の記載内容が不明確、抽象的なものがあり、評価を適切に行うことが難しい。その上、実際の評価実務において、限られた情報及び限られた人員と期間の中で、適切に評価を行うことは、評価担当者に負担がかかり「評価疲れ」が指摘されている。評価の質を担保する上で、評価に必要な要素を漏れなく、過不足なく、評価を行う際に適切に織り込むことは不可欠である。本稿は、こうした独立行政法人制度における評価の理念と現実を踏まえ、独立行政法人制度における実績評価のさらなる適正化を目指し、実績評価モデルを提案する。本モデルを活用することにより、評価担当者は、評価で不足している要素を検出でき、適切な評価実施に貢献するものと考えられる。

キーワード

独立行政法人評価、森田モデル、評価品質、構成要素

1. はじめに

本稿は、独立行政法人制度の評価に係る実績評価モデルの構築を目的とする。ここでいう実績評価モデルは、複雑な法人活動の実態を本質的に表現し、第一義的に独立行政法人の評価に不足している情報の検出を目指すものである。

独立行政法人制度は、中央省庁等改革の柱の一つとして、行政改革会議最終報告（平成9年12月）において、導入が提言されたものである。独立行政法人制度の目的は、各府省の行政活動から、政策の実施部門のうち一定の事務・事業を分離し、これを担当する機関に国家行政組織から独立した法人格を与え、業務の質の向上や

活性化、効率性の向上、自律的な運営、透明性の向上を図ることにある。

独立行政法人制度に類似する制度には、国立大学法人制度、地方独立行政法人制度がある。本稿では、独立行政法人通則法（平成11年法律103号、以下、「通則法」という。）の下に包摂される独立行政法人のみを研究対象とする。

独立行政法人は、法文上、「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直接に実施する必要のないもののうち、民間の主体に委ねた場合には必ずしも実施されないおそれがあるもの又は一の主体に独占して行わせることが必要であるものを効

果的かつ効率的に行わせるため、中期目標管理法人、国立研究開発法人又は行政執行法人として、この法律及び個別法の定めるところにより設立される法人をいう。」(通則法第2条)と定義されている。

平成13年4月、独立行政法人は57法人でスタートした。これらに加えて、平成15年10月以降は、「特殊法人等整理合理化計画」に基づき特殊法人等が独立行政法人に移行した。平成17年10月に、独立行政法人数は113法人までに達した。その後、独立行政法人の統廃合等が行われ、平成30年4月現在、87法人に至っている。

独立行政法人制度は制度導入から、10年超が経過し様々な問題が顕在化してきた。そのため、抜本的な見直しが行われ平成27年4月より、装いを新たにした独立行政法人制度が再スタートした。この制度改正の法的根拠は、平成26年の第186回通常国会に提出され可決された「独立行政法人通則法の一部を改正する法律等」であった。

本稿にとって重要な改正通則法下の新たな独立行政法人制度のポイントは以下の2点である。

第1に、各府省に置かれた独立行政法人評価委員会が評価を行う仕組みから、主務大臣が法人の業務実績の評価を行う仕組みへと変更されたことである。このことにより、従前の各府省に置かれていた独立行政法人評価委員会は廃止された。本稿では、この主務大臣の下で行われることになった新たな業務実績の評価を「実績評価」と称する。

第2に、主務大臣は、法人の業務等に係る国民への説明責任を果たしつつ法人の政策実施機能を最大化するとの観点から、適切な目標を定め、適切に評価を実施する必要があるとされたことである。このことに関連し、総務省は、「独立行政法人の目標の策定に関する指針」及び「独立行政法人の評価に関する指針」(平成26年9月2日総務大臣決定)を策定した。

これらの指針により、評価区分についてはB評定を標準とした5段階の評語を付すことが原則とされた。また、独立行政法人の目標については、その指標の記載例を示すなど、府省横断的、政府統一的に目標策定及び実績評価の方法等が策定された。

こうして生み出される評価結果は、主務大臣の下でのPDCAサイクルを十分に機能させるために用いられるだけではなく、法人の組織・事業の見直しや改廃、国の政策への反映に活用される。その際には、評価結果の品質の確保が重要となる。

すなわち、評価結果は、独立行政法人の業務の実績について、適切な評価を実施したうえで提供されるものでなければならない。

そのためには、当該業務に係る目標の達成度を適切かつ客観的に測定できることが求められる。

しかし、現実には、目標の記載内容が不明確、抽象的であり、評価を適切に行うことが難しい場合もある。その上、実際の実績評価においては、限られた情報及び限られた人員と期間の中で、適切に実績評価を実施することは負担がかり「評価疲れ」が指摘されている。

したがって、評価の質を担保する上で、評価に必要な要素を漏れなく、過不足なく、評価を行う際に適切に織り込むことが不可欠であり、本稿では、その処方箋を提案する。

本稿の構成は下記のとおりである。まず、第2章において、関連研究をレビューする。第3章では、独立行政法人における実績評価の構造を踏まえた上で、実績評価モデルを構築する。第4章では、実績評価モデルに独立行政法人の評価事例を適用し、必要な概念が含まれているかを検証する。最後に、本稿の結論と今後の展望について述べる。

2. 関連研究

まず、独立行政法人制度における評価モデルを構築するにあたり、独立行政法人に関する研究と評価モデルに関する研究をレビューする。

2.1 独立行政法人に関する研究

これまで、独立行政法人に関する研究の多くは、制度論に主眼が置かれてきた。例えば、多賀谷(1998)、岡本(2001)、岡本(2008)などがそれに該当する。これらの一群の研究は、独

立行政法人制度の創設の趣旨から始まり、経緯、制度設計の解説などが大半を占めている。しかし、個別具体的な評価の内容に入った研究は見られない。森田（2015）は、具体的な評価の内容に立ち入って研究を進めたものである。ここでは、独立行政法人の5段階の評価結果を基に、その評定についての考察が行われている。しかし、同研究も評価の構造までは分析が及んでいない。

2.2 評価モデルに関する研究

社会科学におけるモデルについて、経営学者であるAckoff(1962)は、「注目している現象を記述した現実の理想化された表現」と定義している。つまり、モデルとは、ある対象の理解を深めるために事象を取捨し、それを論理的に記述して表現したものである。ここからいえば、何らかのモデルと比較することで、現実の情報の過不足が明確となる。

本稿では、こうした観点から実績評価モデルを構築し、現実の評価実務との比較を行う。すなわち、実績評価モデルを構築することが、的確な評価結果を得るための必要な第一歩であると考える。

評価に関するモデルについては、行政活動の一連の流れを捉える際に、ロジック・モデルが利用されている。また、分野を問わない評価モデルとしてMorita（2015）モデル（以下「森田モデル」という。）がある。以下では、ロジック・モデルと森田モデルについてレビューを行う。

2.2.1 ロジック・モデル

ロジック・モデルは、1970年代に米国の政策シンクタンクである、アーバンインスティテューートのWholeyらによって提唱されたものである。Wholey(1979)は、ロジック・モデルについて、資源のインパクト活動、アウトカムないしは、インパクトの関連を論理的に示すものであるとしている。その背景には、政策実施研究における蓄積があった。

また、ロジック・モデルは、W.K. Kellogg Foundation (1998)が発行した「ケロッグ財団評価ハンドブック」にプログラムロジックモデルと

して紹介されている。W.K. Kellogg Foundation (2004)では、ロジック・モデルの目的について、関係者にロード・マップを示すことにあるとしている。すなわち、「もし～ならば、こうなるだろう」という仮説のもと、資源・インプット、活動（アクティビティ）、直接の結果（アウトプット）、成果（アウトカム）、活動の結果生じた社会的、環境的な成果（インパクト）とを繋ぎ合わせて提示するものである。

ロジック・モデル（W.K. Kellogg Foundation 2004）は、図1のとおり事業が成果を上げるために、必要な要素を体系的に図式化したものである。この各要素の内容は、下記のとおりである。

(1)資源・インプット

プログラム運営には、何らかの資源が必要である。

(2)アクティビティ

もし、当該資源を活用できれば、計画された活動の実現が可能である。

(3)アウトプット

もし、計画した活動が実施されれば、予定した量の製品やサービスの供給が可能である。

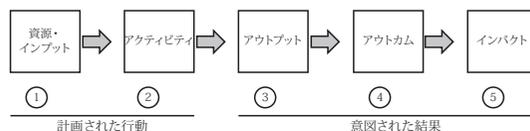
(4)アウトカム

もし、予定した活動が実施されれば、参加者は何らかの利益を得る。

(5)インパクト

もし、参加者が利益を得られれば、組織・地域・システムに何らかの変化が起きる。

図1 基本ロジック・モデル



(出所) W.K.Kellogg Foundation (2004:3) をもとに筆者翻訳

上記のようなW.K.Kellogg Foundationだけでなく、米国ワシントンD.C.に本部がある非営利組織であるInnovation Network,Inc.のロジック・モデルもある。このInnovation Network,Inc.(2010)のロジック・モデルワークブックによれば、ロジック・モデルとは、上位の目標（ミッション）か

ら事業（プログラム）の活動レベルまでの論理的繋がりを視覚的に把握するためのツールであるとされている。

2.2.2 森田モデル

評価者と評価対象との関係について捉えた「森田モデル」がある（Morita 2015）。森田モデルは、プロジェクトを評価する場合の基礎研究として「評価とは何か」という部分に焦点を当て検討し、評価についての網羅的な評価モデルを構築したものとされている（図2）。この森田モデルは、評価者、評価対象、評価の視点、基準の4つから成り立っている。評価者は、評価の視点（品質、コスト、納期）、基準を用いて、評価対象を評価することになる。

この評価者には、評価対象の業務や事業を所管する組織以外の外部組織や第三者、利用者、納税者といった人が含まれる。また、グループ企業であれば親会社が、評価者になる場合がある。さらに、評価者と評価対象が同じ部門や企業といった内部組織の場合もある。Morita（2015）は、評価を受ける評価対象は、ヒト・モノ・カネ・情報からなるとしている。

評価者は、評価の視点と評価を行う際の基準に基づいて、評価対象を評価している。しかし、森田モデルでは、評価についての大枠の関係をモデル化しているものの、評価の視点の提示に

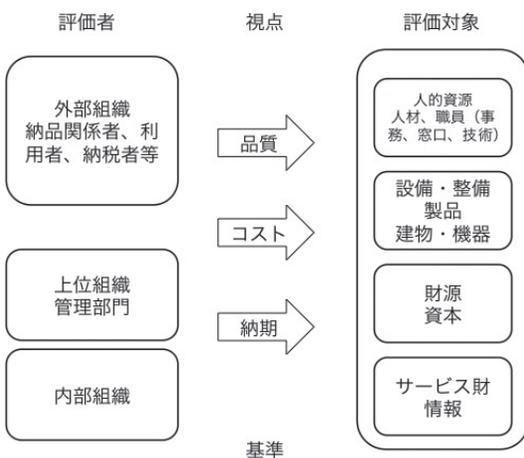
留まっている。同モデルでは、評価を行う際に必要となる評価基準の構成要素について、具体的に明示されていない。

3. 実績評価モデルの検討

関連研究より評価において、評価者、評価対象、評価の視点があり、さらに明確な評価基準が必要であることがわかった。すなわち、評価は、評価目的に照らして、誰が評価を行うのか、何を評価対象とするのかが設定される。この評価対象が決まれば、評価基準を用いて評価がされることになる。したがって、評価基準が、評価を適切に実施する上で重要な要素であり、その基準が不適切なものであれば、当然、正しい評価結果を得ることはできない。また、評価基準が曖昧な場合、評価者によって属人的な評価となり評価結果のレベルに、ばらつきが生じる。これらのことは、評価における信頼性、客観性を損ねる要因となる可能性がある。そこで、評価基準のモデルを含んだ実績評価モデルを構築する必要がある。

本稿では、こうした一連の評価全体を表現したモデルを作成し、このモデルを「実績評価モデル」と呼ぶ。実績評価モデルを構築するにあたり、評価者、評価対象、評価基準における各要素について以下、検討を行う。

図2 森田モデル



(出所) Morita (2015:391) をもとに筆者翻訳

3.1 評価者の構成要素

評価者には、評価対象との関係において、ここでいう内部と外部の評価者に分類できる。内部には、評価者と評価対象が同じ組織・部門の場合があると考えられる。さらに、外部には、外部の専門家や、第三者機関など別の組織の場合が考えられる。

3.2 評価対象の構成要素

評価の目的によって、評価対象が異なり、それに伴い評価基準も変わる。そこで、ロジック・モデルの分析の流れにしたがって、その要素と森田モデルを踏まえて、以下、評価対象の構成要素について検討を行う。

(1)インプット

事業の活動を実施するために、資源をインプットする必要がある。インプットには、業務を行う職員、人材、専門技術、オフィスや施設、資材、材料、事業費、予算、各種の経費、業務を行う上で必要となる情報等といったものが考えられる。

評価対象の構成要素として森田モデルでは、ヒト・モノ・カネ・情報があるとしている。すなわち、主要な経営資源として、ヒト・モノ・カネといった有形財産と情報といった無形財産とがある。

(2)アクティビティ

アクティビティとは、事業を通じて提供するサービス等を産み出すための具体的な事業活動である。インプットとして投入された経営資源を利用して、業務や事業といった活動が行われる。よって、ヒトに関する活動、モノに関する活動、カネに関する活動、情報に関する活動として、アクティビティは表現できると考えられる。各活動は組織単独で行うこともあれば、他の組織と協力をして行うこともある。例えば、業務の一部をアウトソースすることが挙げられる。したがって、協力者・協力組織といった要素が考えられる。

(3)アウトプット

アウトプットとは、投入した行政資源を利用して活動をしたことによって、直接産み出された結果である。例えば、道路等の整備延長、研修の実施回数などといったサービスの提供などがある。

(4)アウトカム

活動の直接の結果（アウトプット）がもたらす、達成しようとする成果である。例えば、国民、受益者、経済・社会の状況等といった国民の生活及び社会経済に及ぼす達成しようとする影響や変化がある。ロジック・モデルには、上記の要素の他にインパクトがある。インパクトとは、より長期的なアウトカムのことを指す。具体的には、活動の成果として、数年後に生じ

ることが予想される組織、地域、制度レベルの変化や、予想外の変化など長期的な成果がここに含まれる。したがって、インパクトを評価するためには、評価の射程を広範囲に設定しなければならない。しかし、評価に投入できる資源には制約があるため、実績評価モデルを構築するにあたり、インパクトは考慮しない。

3.3 評価基準の構成要素

先述のロジック・モデルや森田モデルは、それぞれ重要な視点を示している一方で、どのように計測するかについて、具体的には述べられていない。そこで、次に評価基準の構成要素について検討を行う。

3.3.1 尺度

評価を行うためには、その対象物を何らかの視点によって計測する必要がある。計測できない限り、評価することはできない。例えば、鉄球は大きさ（長さ）、重さ、硬さなどの視点で計測できる。この計測のための視点は複数存在する。ある視点で、計測のための単位を決めているものを尺度と呼ぶ。鉄球のような、質量のある物体の場合は、SI（国際単位系：International System of Units）が尺度として用いられる。

一方、組織活動などのプロセスについての尺度を明確にすることは難しい。しかし、組織活動を計測する方法は、経験則的に単位として、人月や、金額といった量で表記され、計測可能性を実現する。また、サービスといったものについては、利用者数や、サービスを受けた側の満足度といった尺度も考えられる。しかしながら、これらは必ずしもSIのように国際的に標準化されているわけではない。このことから、評価者もしくは、サービス実施者が尺度を定義する必要がある。

3.3.2 基準値と比較対象

尺度を評価に組み替えるためには、基準値が必要となる。基準値が記述されていない場合、評価者は、評価対象について測定した値の良し悪しの判断を行うことができない。そのため、実績値が高いのか低いのか、望ましいのか望ま

しくないのかを判断するために、基準値が必要である。

さらに、基準を作るためには、比較対象がなければならない。比較をすることにより、客観的に評価が可能となるからである。比較対象としては、前年度比のように、自己の過去との比較をした時系列比較や、ベンチマーキングのような、他の組織と比較する同等他者の比較がある。このように、評価基準は、尺度、基準値、比較対象から構成される。

3.4 実績評価モデルの構築

評価モデルに必要な要素について、検討をしてきた。結果を踏まえて、評価全体のモデルを構築し、図3に示す。このモデルを実績評価モデルと呼ぶこととする。

実績評価モデルは、①評価者（内部、外部）、②評価対象（インプット、アクティビティ、アウトプット、アウトカム）、③評価基準（尺度、基準値、比較対象）の3つの要素からなる。同モデルは、独立行政法人の活動が、評価される過程を可視化している。Lave and March(1975=1991)によれば、モデルの適用範囲が広いほど、その

モデルは、よいモデルであり、様々なインプリケーションを考えうるとし、また、美しいモデルは単純であるとも述べている。本稿でも、複雑になることを避け、シンプルに本質を表現することを心がけて作成した。

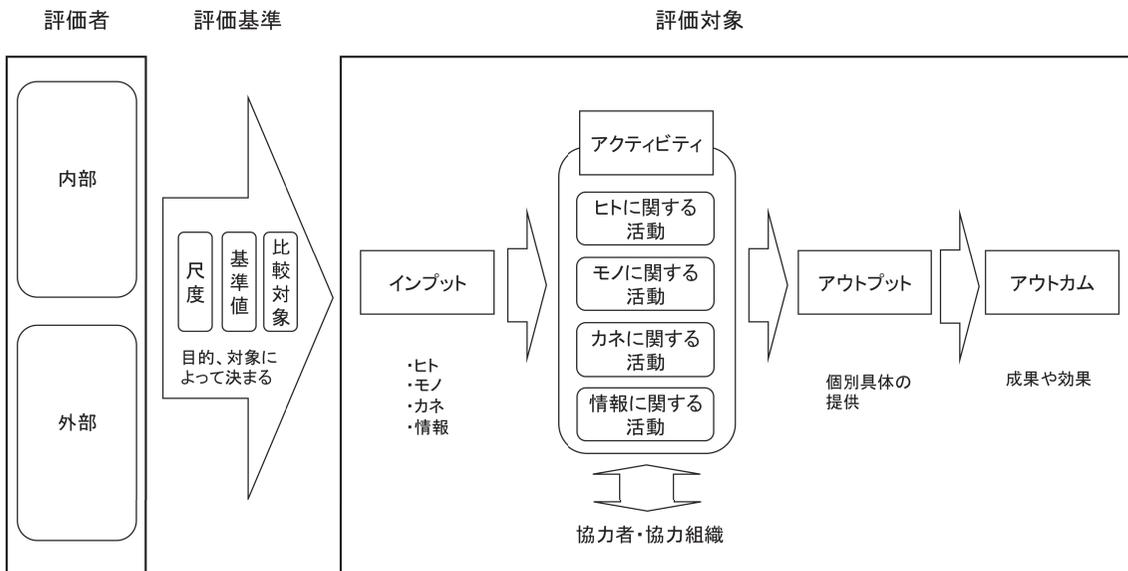
4. 実績評価モデルへの適用

本稿は、独立行政法人評価における評価基準を含んだ評価全体のモデルを構築することが目的である。そのため、実績評価モデルを構成する3つの要素（評価者、評価対象、評価基準）に、独立行政法人評価の事例を適用することにより、実績評価モデルに必要な概念が含まれているか確認を行う必要がある。

4.1 評価者の構成要素への適用

冒頭に述べたとおり、平成26年の制度改正により、評価主体は各府省に設定された第三者委員会である独立行政法人評価委員会から、主務大臣に変更がされている。また、総務省に置かれた外部有識者から構成される独立行政法人評

図3 実績評価モデル



(出所) 筆者作成

価制度委員会は、主務大臣が行う実績評価に対して、必要に応じて意見を述べるスキームとなっている。

上記を分解した各要素が実績評価モデルのどの構成要素に相当するかを確認すると、内部の評価者には、当該独立行政法人の主務大臣が該当する。また、外部の評価者には、第三者機関としての独立行政法人評価制度委員会が該当する。なお、ここでいう評価者は、評価主体において、実際に評価の実務を行う評価担当者も含み幅広く捉えている。

4.2 評価対象の構成要素への適用

4.2.1 分析対象の概要

実績評価モデルにおいて、評価対象の構成要素が、抽出されているのか確認を行うために、評価事例を評価対象の構成要素に適用する。そのために、本稿では、評価事例を下記の方法で抽出した。独立行政法人は、業務の特性により、①中期目標管理法、②行政執行法人、③国立研究開発法人の3つに分類されている。

そこで、平成27年度の年度評価の結果をまとめた評価書より、各分類より代表的な3法人として①独立行政法人国際観光振興機構（事例1）、②独立行政法人国立公文書館（事例2）、③国立研究開発法人防災科学技術研究所（事例3）を抽出した（以下、「法人名称」を一部省略し表記する）。

年度評価とは、評価対象年度以降の業務運営の改善に資するために、法人の業務実績全体について、統合的な評定を行うものである。この評価書を基に、実績評価モデルの評価対象について適用を行った。その結果をまとめたものが、表1である。

表1の各項目の内容と、その作成手順について説明する。表1の項目「事業概要」については、評価書の「1.当事業及び事業に関する基本情報」の内容を抽出した。業務の特性によって、独立行政法人は3つに分類される。それぞれの法人のタイプに応じて、目標期間も異なっている。具体的には、①中期目標管理法では、「中期目標」、②行政執行法人では、「年度目標」、③国立研究開発法人では、「中長期目標」となっている。

上記の①～③は、いずれにおいても、各目標に基づき評価が行われる。そのため、各目標は評価基準の一部を構成する。このことを踏まえ、「目標」の項目には、それぞれの分類に該当する目標から内容を抽出した。

次に、事例1と2は、評価書の原文にある「法人の業務実績・自己評価」の項目、事例3は、「法人の業務実績等・自己評価、主な業務実績」を読むことにより、その内容を「アクティビティ」、「アウトプット」、「アウトカム」の項目に整理分類した。表1では、これらの要素を横軸に表記している。

なお、実績評価モデルの要素である「インプット」については、評価書に該当する記載がある法人とない法人があるため、表1については「インプット」を割愛する。但し、後述の図4～6については、法人が活動するための資源をインプットすることにより、アクティビティを行えるため、インプットを表記している。

4.2.2 結果

評価事例について分析をした結果は、以下の通りである。

(1)事例1: 国際観光振興機構

国際観光振興機構は、平成15年に設立した国土交通省所管の法人であり、中期目標管理法に分類されている。国際観光振興機構は、マーケティングの高度化を図りつつ、現地目線に立った効果的なプロモーション活動を実施し、インバウンドの更なる拡大のため、地方への誘客、質の向上に向けた取り組みを強力に推進することを主な業務としている。

業務活動の一例として、表1のA-3より、ツーリスト・インフォメーション・センター（TIC）の運営がある。この業務活動は国内受入体制整備支援業務において、観光案内所の整備支援業務として行われている。この業務活動について、実績評価モデルに適用すると図4となる。

表1のA-3では、「運営にあたっては、ウェブを活用した情報提供コンテンツを創設するとともに、TICにおいて対面、電話又はウェブによって情報提供を行う機会を増加させる。」ことを目標

表1 評価対象の構成要素への適用

国際観光振興機構					
No.	事業概要	目標	アクティビティ	アウトカム	
A-1	訪日プロモーション業務 ②訪日外国人旅行者誘致のための業務	観光庁と共同で、市場ごとに海外の市場動向等の情報に基づく効率的・効果的な訪日プロモーション方針を策定・公表し、地方自治体や民間事業者が独自に取り組む事業・活動との連携を図りやすい環境を整えるとともに、訪日プロモーション事業の実施主体として、オールジャパン体制で国としての日本の観光魅力を発信し、訪日に結びつける取組を強力に実施する。	市場別訪日プロモーション方針に基づく事業実施	<ul style="list-style-type: none"> 観光庁と機構が共同で策定し、マーケティング戦略本部で決定した、市場別の訪日プロモーション方針に基づき、事業を実施した。訪日プロモーション方針は、重点20市場及び欧州潜在市場の合計21市場別に策定し、5月に公表した。 機構内に、ビジット・ジャパン事業実施本部を4月に新設し、事業の進捗を管理するとともに、事業計画から実施結果・成果の分析を行う等、PDCAの確立を図った。 	<ul style="list-style-type: none"> オールジャパン体制での取組を促進するとともに、海外の旅行会社等の招請、企画コンサルティング、共同キャンペーン等の実施を通じて、2015年に過去最多を記録した訪日外国人旅行者数(1,974万人)の大幅な増加(47.1%増)に貢献した。
A-2	国内受入体制整備支援業務 ①観光案内所の整備支援業務	支援の実施に当たっては、認定案内所からの評価を定量的に把握し、相当程度の評価を得るものとする。機構が運営するツーリスト・インフォメーション・センター(以下「TIC」という。)については、機構がこれまでに培ってきたネットワークや手法に基づき、民間のノウハウを取り入れた効果的な運営を行うとともに案内所支援業務の中核として全国の案内所との連携を強化する。	外国人観光案内所の認定・支援	<ul style="list-style-type: none"> 観光案内所認定に係る電子申請システムを新たに構築し、認定案内所の管理データベースと統合することにより業務の効率化を図った。 	<ul style="list-style-type: none"> TICは、民間のノウハウを取り入れた運営を行い、全国各地の外国人観光案内所へのノウハウ提供、問い合わせ対応等を通じて認定外国人観光案内所のネットワーク拡大及び質のレベルアップに貢献した。
A-3	国内受入体制整備支援業務 ①観光案内所の整備支援業務	運営にあたっては、ウェブを活用した情報提供コンテンツを創設するとともに、TICにおいて対面、電話又はウェブによって情報提供を行う機会を増加させる。	ツーリスト・インフォメーション・センター(TIC)の運営	<ul style="list-style-type: none"> ウェブを活用した情報提供コンテンツの拡充。 大規模災害等発生時の緊急時の電話対応24時間化。 	<ul style="list-style-type: none"> ウェブによる情報提供件数は、目標値の25万件を大きく上回る39.5万件となり、ウェブによる情報ニーズの高まりに対応した。 平成26年度に整備した震度6以上(東京23区内は震度5強以上)の地震発生時におけるTICの24時間電話対応体制について、5月に発生した小笠原諸島沖地震(M8.5)時に初めて稼働させ、夜間電話対応を実施した。
国立公文書館					
No.	事業概要	目標	アクティビティ	アウトカム	
B-1	②歴史公文書等の適切な移管及び保存に向けた行政文書等の管理に関する適切な措置	ii) 行政機関及び独立行政法人等における歴史公文書等の選別等に係る適切な判断を支援するため、必要に応じて専門的技術的助言を行うこと。	専門的技術的助言	<ul style="list-style-type: none"> 約321万件の行政文書ファイル等に対して保存期間満了時の措置の適否に係る専門的技術的助言を実施した。確認作業の効率化等により、件数が前年度(約267万件)の約20%増となった。なお、平成27年度満了文書のうち、年度内に専門的技術的助言ができなかった残余約86万件については、引き続き精査作業を進め、平成28年度に助言を行うこととしている。 約132万件の行政文書ファイル等に対して、廃棄に係る協議に関する専門的技術的助言を実施した。 平成27年度末までに保存期間が満了する法人文書ファイル等について、館への移管に関する意向調査を実施した。当該調査に対する独立行政法人等からの回答を受け、移管希望のあった法人に対して、移管基準への該当性等についての調査・照会等を行い、専門的技術的助言等を行った。 	<ul style="list-style-type: none"> 事業計画上の数値目標として、行政文書ファイル等のレコードスケジュールの適否に係る専門的技術的助言を前年度比約10%増にあたる約300万件以上と定め、効率的な確認作業を進めた結果、前年度比約20%増の約321万件の助言を行った。
B-2	(4)歴史公文書等の利用の促進その他の措置 ②デジタルアーカイブの運用及び充実	②デジタルアーカイブの運用及び充実 i) 館の保存する特定歴史公文書等を広く一般の利用に供するため、インターネットにより所蔵資料を検索し、閲覧できるデジタルアーカイブを推進すること。また、外部の意見を聴取し、デジタルアーカイブの充実を図ること。	②デジタルアーカイブの運用及び充実	<ul style="list-style-type: none"> 紙資料から直接デジタル化する方法により、約210万コマのデジタル化を行い、館デジタルアーカイブへ登録した。これにより、既に公開している約1,314万コマと合わせ、約1,524万コマのデジタル画像を館デジタルアーカイブに登録の上、インターネットでの提供を開始した。 本年度も複製作成計画に基づき、計画的なデジタル化を実施した結果、特定歴史公文書等のデジタル画像の作成率は、12.9%となり、デジタルアーカイブの充実を図った。 デジタル化に当たっては、ウェブアンケートを実施し、内閣文庫資料のデジタル化に係るニーズ等を把握した。 	<ul style="list-style-type: none"> システムの効率化が図られるとともに、パソコンだけでなく、タブレット端末やスマートフォンにも対応し、より使い易く、デジタル画像等のダウンロードも可能となるなど、デジタル情報資源の提供の仕組みの強化等による更なる国民の利便性の向上を図った。
防災科学技術研究所					
No.	事業概要	目標	アクティビティ	アウトカム	
C-1	災害リスク情報に基づく社会防災システム研究	②特に地震災害に関しては、全国を対象とした地震ハザード・リスク評価手法の高度化及びそれら成果の地域への展開に取り組むとともに、ハザード・リスク評価の基盤となる地下構造に関する情報や活断層に関する情報の整備を行う。また各種情報を公開し、利活用を促進するためのシステム開発を実施する。津波災害に関しても、全国を対象とした津波ハザード評価手法を開発する。	b) 全国津波ハザード評価手法の開発	<ul style="list-style-type: none"> 津波ハザード評価については、南海トラフ、相模トラフ、日本海溝で発生する地震に対する津波ハザードの評価を実施。 津波ハザード情報の利活用に関する検討を実施し報告書を取りまとめた。 	<ul style="list-style-type: none"> 地震調査研究推進本部の津波評価部会に資料を提出し、国の施策に貢献。 外部資金による事業と連携し日本海の地震による津波に対しての波源モデルの検討を進め、日本周辺での主要な地震による津波ハザード評価ができる見込みとなった。
C-2	災害リスク情報に基づく社会防災システム研究	③自然災害に関するハザード・リスクを評価する技術については、国際的な普及に向けて先導することを旨とし、国際機関や海外の研究機関と連携しつつ進める。	c) 各種自然災害リスク評価システムの研究開発	<ul style="list-style-type: none"> 地震以外の各種災害に対するリスク評価については、自然災害事例データベースの構築を進めるとともに、地すべり地形分布図を完成させた。 	<ul style="list-style-type: none"> 日本全国どの場所でも、人が住む地域の自然災害の脆弱性を一目で理解できる仕組みを通して防災力向上。
C-3	災害リスク情報に基づく社会防災システム研究	③自然災害に関するハザード・リスクを評価する技術については、国際的な普及に向けて先導することを旨とし、国際機関や海外の研究機関と連携しつつ進める。	d) ハザード・リスク評価の国際展開	<ul style="list-style-type: none"> 外部資金による事業と連携して風水害ハザード・リスク評価の研究を進めるとともに、雪氷災害等に関しては所内の他のプロジェクトと連携して研究を進めた。 ハザード・リスク評価の国際展開に関しては、アジア地域各国との共同研究を継続するとともに、国際NPO法人GEMの活動に積極的に関与するなど、我が国で培ってきた各種知見を国際的に広める。 	<ul style="list-style-type: none"> アジア地域での地震ハザード評価に関する取組を強化。

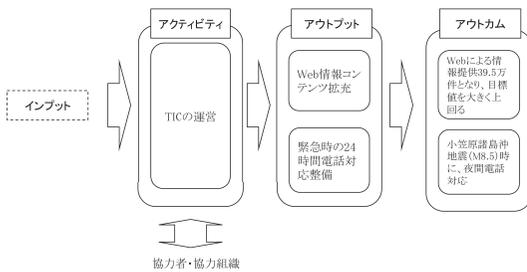
(出所) 各独立行政法人の評価書の内容をもとに筆者作成

としていると記載されている。

評価書の「主要なインプット情報」の欄は、空欄となっているが、当該目標を達成するために、「ツーリスト・インフォメーション・センター（TIC）の運営」というアクティビティを行ったとされている。そのアウトプットとして、「ウェブを活用した情報提供コンテンツの拡充」、「大規模災害等発生時の緊急時の電話対応24時間化」を行ったということが記述されている。

さらに、表1を見ていくと「ウェブによる情報提供件数は、目標値の25万件を大きく上回る39.5万件となり、ウェブによる情報ニーズの高まりに対応した。」、「平成26年度に整備した震度6以上（東京23区内は震度5強以上）の地震発生時におけるTICの24時間電話対応体制について、5月に発生した小笠原諸島沖地震（M8.5）時に初めて稼働させ、夜間電話対応を実施した。」といったアウトカムが記述されている。

図4 実績評価モデル適用事例1



(出所) 筆者作成

(2)事例2: 国立公文書館

国立公文書館は、平成13年に設立した内閣府所管の法人であり、行政執行法人に分類されている。国立公文書館は、国の各機関から受け入れた歴史資料として重要な公文書等を将来にわたり確実に保存し、これらを閲覧・展示などを通じ広く国民の利用に供することを主な任務としている法人である。表1のB-2では、歴史公文書等の利用の促進その他の措置の一つとして、デジタルアーカイブの運用及び充実が行われていることが示されている。この業務活動について、実績評価モデルを適用すると図5となる。

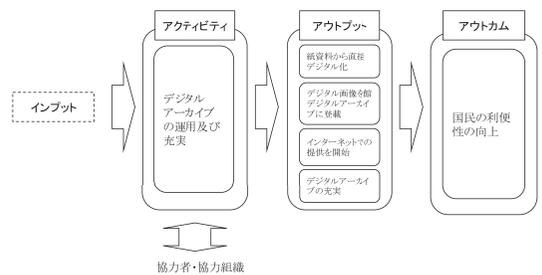
ここでの目標は、「デジタルアーカイブの運用及び充実」とされるとともに、「館の保存する特定歴史公文書等を広く一般の利用に供するため、インターネットにより所蔵資料を検索し、閲覧できるデジタルアーカイブを推進すること。また、外部の意見を聴取し、デジタルアーカイブの充実を図ること。」とされている。

評価書の「主要なインプット情報」の欄は、空欄となっているが、当該目標を達成するために、「デジタルアーカイブの運用及び充実」というアクティビティが設定されている。

さらに、その結果については、「紙資料から直接デジタル化する方法により、約210万コマのデジタル化を行い、館デジタルアーカイブへ登録した。これにより、既に公開している約1,314万コマと合わせ、約1,524万コマのデジタル画像を館デジタルアーカイブに登録の上、インターネットでの提供を開始した。」、「本年度も複製物作成計画に基づき、計画的なデジタル化を実施した結果、特定歴史公文書等のデジタル画像の作成率は、12.9%となり、デジタルアーカイブの充実を図った。」、「デジタル化に当たっては、ウェブアンケートを実施し、内閣文庫資料のデジタル化に係るニーズ等を把握した。」といったアウトプットが得られたと述べられている。

また、「システムの効率化が図られるとともに、パソコンだけでなく、タブレット端末やスマートフォンにも対応し、より使い易く、デジタル画像等のダウンロードも可能となるなど、デジタル情報資源の提供の仕組みの強化等による更なる国民の利便性の向上を図った。」というアウトカムも示されている。

図5 実績評価モデル適用事例2



(出所) 筆者作成

(3)事例3: 防災科学技術研究所

防災科学技術研究所は、平成13年に設立した文部科学省所管の法人であり、国立研究開発法人に分類されている。防災科学技術研究所は、災害から人命を守り、災害に強い社会を実現するために、防災科学技術に関する基礎研究及び基盤的研究開発等の業務をしている。

表1のC-3では、「災害リスク情報に基づく社会防災システム研究」が掲げられている。この業務活動について、実績評価モデルに適用するものが図6である。

C-3の目標は、「自然災害に関するハザード・リスクを評価する技術については、国際的な普及に向けて先導することを目指し、国際機関や海外の研究機関と連携しつつ進める。」こととされている。評価書の「主要なインプット情報」の欄に「予算額と従事人員数」の記載があり、当該目標を達成するために、「ハザード・リスク評価の国際展開」というアクティビティが行われていると説明されている。

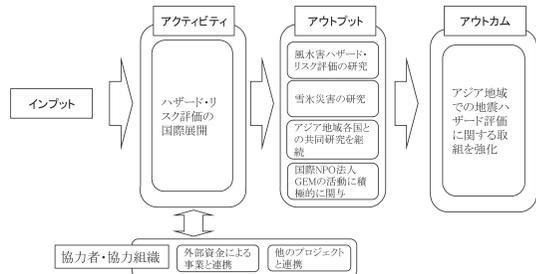
アウトプットについては、「外部資金による事業と連携して風水害ハザード・リスク評価の研究を進めるとともに、雪氷災害等に関しては所内の他のプロジェクトと連携して研究を進めた。」「ハザード・リスク評価の国際展開に関して、アジア地域各国との共同研究を継続するとともに、国際NPO法人GEMの活動に積極的に関与するなど、我が国で培ってきた各種知見を国際的に広める。」とされている。

また、前述のとおり「外部資金による事業と連携」や、「他のプロジェクトと連携」といった、協力者・協力組織という構成要素も、ここには含まれている。アウトカムとしては、「アジア地域での地震ハザード評価に関する取組を強化」とある。

以上より、事例1から3を実績評価モデルに適用した結果、アクティビティ、アウトプット、アウトカムの各要素は、評価対象の構成要素に該当することが確認できた。そして、アクティビティを通じて、情報の提供や研究成果の公開といったサービスの提供があることも判明した。その際に、例えば、共同研究といった協力者・協力組織と連携して活動を行っていることも伺

えた。さらに、インプットについて評価書に記載がないものがあることが把握できた。

図6 実績評価モデル適用事例3



(出所) 筆者作成

4.3 評価基準の構成要素への適用

4.3.1 分析概要

実績評価モデルの評価基準の構成要素が抽出されているのか確認するために、前述と同様の事例を用いる。その結果をまとめたものが、表2である。

表2の各項目の内容と、その作成手順について説明をする。その手順は下記の(1)～(5)である。以上を踏まえ、適用事例の結果を表2にまとめた。

- (1)表2の項目「事業概要」については、評価書の「1.当事業及び事業に関する基本情報」の内容を抽出した。
- (2)「評価基準」は、評価書の原文にある「主な評価指標」から、事例1と2は、〈主な評価指標〉より、事例3は、〈主な評価軸（評価の視点）、指標等〉より抽出した。さらに、抜粋した原文に、1)、2)・・・と番号をふり、わかりやすくした。
- (3)「尺度」は、(2)の「評価基準」から尺度となる原文よりまとめて抽出した。さらに、但書きで追加として考えられる尺度を記載した。
- (4)「基準値」は、(3)を基に基準値について、明示されている場合には、その基準値を記載し、基準値が明示されていない場合に「記載なし」とした。
- (5)「比較対象」は、(3)を基に比較対象について、明示されている場合には、その比較対象を記載し、比較対象が明示されていない場合には、「記載なし」とした。

表2 評価基準の構成要素への適用

法人名	No.	事業概要	評価基準	尺度	基準値	比較対象
国際観光振興機構	a-1	(1) 訪日プロモーション業務①海外事務所を活用した市場動向の収集・調査・分析・提供	<主な定量的指標> 1) 事業パートナーに対し、海外事務所員や本部職員による面談方式の個別コンサルティングを年間延べ3,000件以上実施し、インバウンド関係者の育成とインバウンドビジネスの活性化を目指す。 2) 事業パートナーに対する調査において、機構からの情報提供が、4段階評価で上位2つの評価を得る割合が7割以上となることを目指す。 <その他の指標> なし	1) 個別コンサルティングを年間延べ3,000件以上実施 2) 機構からの情報提供が、4段階評価で上位2つの評価を得る割合が7割以上	1) 3,000件以上 2) 4段階評価で上位2つの評価を得る割合が7割以上	1) 年間延べ 2) 記載なし
	a-2	(2) 業務運営の効率化	<主な定量的指標> なし <その他の指標> なし <評価の視点> 1) 運営費交付金を充当して行う業務については、一般管理費及び業務経費のうち効率化対象の合計について、毎年度平均で前年度比1.25%以上の効率化を行っているか。 2) 公正性・透明性を確保しつつ、自律的かつ継続的に調達等の合理化に取り組み、業務運営の効率化を図っているか。	1) 効率化対象の合計について、毎年度平均で前年度比1.25%以上の効率化 2) 調達等の合理化に取り組み、業務運営の効率化(但し、調達の合理化率が考えられる。)	1) 1.25%以上 2) 記載なし	1) 前年度比 2) 記載なし
	a-3	(3) 関係機関との連携強化	<主な定量的指標> なし <その他の指標> なし <評価の視点> 1) 在外公館を始めとする関係省庁・政府関係法人、地方公共団体、経済界との連携を強化しているか。 2) 国際協力機構、国際交流基金及び日本貿易振興機構の海外事務所と、海外事務所の共用化又は近接化を実現しているか。 3) 本部事務所については、工程表に基づき、国際交流基金との共用化に向けた準備を適切に行っているか。	1) 連携を強化しているか。(但し、連携先数などが考えられる。 2) 共用化又は近接化を実現しているか。(但し、共用化率や近接件数などが考えられる。 3) 共用化に向けた準備は適切か。(但し、共用化率や近接件数などが考えられる。)	1) 記載なし 2) 記載なし 3) 記載なし	1) 記載なし 2) 記載なし 3) 記載なし
国立公文書館	b-1	(4) 歴史公文書等の利用のための適切な措置①展示等の実施	<主な定量的指標> 1) 展示会等実施回数 <その他の指標> 2) 展示会入場者数 3) 外部の意見を聴取 4) 貸出し	1) 展示会等実施回数 2) 展示会入場者数 3) 外部の意見を聴取(但し、聴取件数などが考えられる。 4) 貸出し(但し、貸出し件数などが考えられる。)	1) 記載なし 2) 記載なし 3) 記載なし 4) 記載なし	1) 記載なし 2) 記載なし 3) 記載なし 4) 記載なし
	b-2	(6) アジア歴史資料データベースの構築及び利用促進	<主な定量的指標> 1) 公開画像数 2) 既公開目録データの遡及点検数 3) ホームページ上のコンテンツの新規公開又は更新 4) 国内外の学会、教育・研究機関における広報活動 5) アジ歴ニュースレターの発行 <その他の指標> 6) 国立公文書館アジア歴史資料センター諮問委員会開催状況 7) 戦後資料の受入れに向けた取組状況 8) 検索精度向上に向けた取組の実施状況 9) リンクによる資料提供機能の拡充の状況 10) アンケート調査等の実施状況 11) アジ歴ニュースレター登録者数	1) 公開画像数 2) 既公開目録データの遡及点検数 3) コンテンツの新規公開又は更新(但し、コンテンツの新規公開数・更新頻度などが考えられる。 4) 広報活動(但し、プレスリリース数などが考えられる。 5) ニュースレターの発行(但し、年間発行回数などが考えられる。 6) 諮問委員会開催状況(但し、諮問委員会開催回数などが考えられる。 7) 受入れに向けた取組状況(但し、資料受入れ件数などが考えられる。 8) 検索精度向上に向けた取組の実施状況(但し、システム改修率などが考えられる。 9) リンクによる資料提供機能の拡充の状況(但し、リンク件数などが考えられる。 10) アンケート調査等の実施状況(但し、アンケート調査件数などが考えられる。 11) ニュースレター登録者数(但し、登録者数などが考えられる。)	1) 記載なし 2) 記載なし 3) 記載なし 4) 記載なし 5) 記載なし 6) 記載なし 7) 記載なし 8) 記載なし 9) 記載なし 10) 記載なし 11) 記載なし	1) 記載なし 2) 記載なし 3) 記載なし 4) 記載なし 5) 記載なし 6) 記載なし 7) 記載なし 8) 記載なし 9) 記載なし 10) 記載なし 11) 記載なし
防災科学技術研究所	c-1	先端的実験施設の整備・共用	【研究環境の充実の観点】 1) 防災科学技術の基盤となる観測・設備等の整備・充実が図られているか。 【イノベーションの観点】 2) 国内外の大学、研究機関、自治体、民間事業者との連携・協力の取組が図られているか。 3) 我が国全体の防災科学技術の水準の向上を図るため、外部の研究開発機関等との共用が進められたか。 <定量的評価> 4) 実大三次元震動破壊実験施設(兵庫県三木市): 25件以上/5年 5) 大型耐震実験施設(茨城県つくば市): 42件以上/5年 6) 大型降雨実験施設(茨城県つくば市): 40件以上/5年 7) 雪氷防災実験施設(山形県新庄市): 110件以上/5年	1) 整備・充実の状況 2) 連携・協力の取組状況(但し、連携・協力件数などが考えられる。 3) 共用の進捗(但し、進捗率などが考えられる。 4) 施設の共用件数: 25件以上/5年 5) 施設の共用件数: 42件以上/5年 6) 施設の共用件数: 40件以上/5年 7) 施設の共用件数: 110件以上/5年	1) 記載なし 2) 記載なし 3) 記載なし 4) 25件以上/5年 5) 42件以上/5年 6) 40件以上/5年 7) 110件以上/5年	1) 記載なし 2) 記載なし 3) 記載なし 4) 年間0件以上 5) 年間0件以上 6) 年間0件以上 7) 年間0件以上
	c-2	研究成果の普及・活用促進及び研究成果の国民への周知	【アウトリーチの観点】 1) 研究・開発の成果・取組の価値を社会に向けて情報発信する取組を推進しているか。 2) 国や地方公共団体、学会、学術誌等で積極的に発表・公開を進めたか。 3) 基礎的地震・火山観測網、Eーディフェンスによって収集されるデータ等の公開に当たっては、より利用しやすいように継続的な改良を進めたか。 4) 報道機関等を通じた情報発信、研究施設的一般公開やホームページによる研究成果の発信、シンポジウムやワークショップの開催などを積極的に進めたか。 <定量的評価> 5) 防災科学技術に関連する査読のある専門誌: 5編/人以上/5年 6) SCI 対称誌等: 240編以上/5年 7) 学会での発表: 30件/人以上/5年・ホームページ(データベースを含む)へのアクセス数: 6,000万件以上/5年 8) シンポジウムやワークショップなどの開催: 100回以上/5年 【イノベーションの観点】 9) 成果の社会実装につながる取組が図られているか。	1) 情報発信の取組の推進(但し、進捗率などが考えられる。 2) 発表・公開の推進(但し、進捗率などが考えられる。 3) 改良の推進(但し、進捗率などが考えられる。 4) 情報発信、研究成果の発信、シンポジウムやワークショップの開催などの推進(但し、進捗率などが考えられる。 5) 防災科学技術に関連する査読のある専門誌数: 5編/人以上/5年 6) SCI 対称誌等掲載数: 240編以上/5年 7) 学会での発表件数: 30件/人以上/5年・ホームページ(データベースを含む)へのアクセス数: 6,000万件以上/5年 8) シンポジウムやワークショップなどの開催回数: 100回以上/5年 9) 取組の状況(但し、社会実装件数などが考えられる。)	1) 記述なし 2) 記述なし 3) 記述なし 4) 記述なし 5) 5編/人以上/5年 6) 240編以上/5年 7) 30件/人以上/5年、6,000万件以上/5年 8) 100回以上/5年 9) 記述なし	1) 記述なし 2) 記述なし 3) 記述なし 4) 記述なし 5) 年間0編/人以上 6) 年間0編以上 7) 年間0件/人以上、年間0件以上 8) 年間0回以上 9) 記述なし

(出所) 各独立行政法人の評価書の内容をもとに筆者作成

4.3.2 結果

実績評価モデルの評価基準の構成要素に評価事例を適用した。その結果、実績評価モデルは、必要な要素を含んでいることが確認できた。さらに、この適用過程において、評価基準の3つの要素である「尺度」、「基準値」、「比較対象」について、一部の要素の記載がないものが検出された。

評価基準の項目数が多いため、適用した図を示すことを省略し、表2に結果をまとめた。以下、表2より4件の指摘を行う。

第1に、表2のa-1より、事業概要の項目「訪日プロモーション業務①海外事務所を活用した市場動向の収集・調査・分析・提供」については、「事業パートナーに対し、海外事務所員や本部職員による面談方式の個別コンサルティングを年間延べ3,000件以上実施し、インバウンド関係者の育成とインバウンドビジネスの活性化を目指す。」という評価基準がある。この尺度は「個別コンサルティングを年間延べ3,000件以上実施」であり、基準値としては「3,000件以上」とあり、比較対象としては「年間延べ」と記載がある。

第2に、a-2より、事業概要の項目「業務運営の効率化」は、「運営費交付金を充当して行う業務については、一般管理費及び業務経費のうち効率化対象の合計について、毎年度平均で前年度比1.25%以上の効率化を行っているか。」という評価基準がある。

この尺度は「効率化対象の合計について、毎年度平均で前年度比1.25%以上の効率化」とあり、基準値は「1.25%以上」と設定されており、比較対象としては、「前年度比」との記載がみられる。

第3に、b-1より、事業概要の項目「歴史公文書等の利用のための適切な措置①展示等の実施」は、「展示会等実施回数」という評価基準がある。この尺度は、「展示会等実施回数」とあるが、基準値や比較対象についての記載がない。

第4に、c-1より、事業概要の項目「先端の実験施設の整備・共用」については、「国内外の大学、研究機関、自治体、民間事業者との連携・協力の取組が図られているか。」という評価基準がある。この尺度は、「連携・協力の取組状況」とあるが、基準値や比較対象についての記載がない。

以上より、評価事例を実績評価モデルの評価基準の構成要素に当てはめた結果、「尺度」、「基準値」、「比較対象」がモデルの要素として抽出されていることが確認できた。さらに、評価基準において、基準値や比較対象を設定できるにも係わらず、それが設定されていないという課題が把握できた。

5. 結論と今後の展望

本稿は、実績評価モデルの構築を通じて、評価の構造及び各要素と要素の関係を明らかにするものであった。同モデルにより、独立行政法人評価にかかるプロセスを体系化することに貢献できるのではないかと考える。

本稿の主張をまとめると以下の通りである。まず、関連研究を踏まえ、評価者、評価対象、評価基準からなる実績評価モデルを構築した。実績評価モデルにおける評価基準は、尺度、基準値、比較対象の3つの要素から構成されているとした。

さらに、実績評価モデルに独立行政法人の評価事例を適用することにより、実績評価モデルが必要な概念要素を含んでいることを確認した。なお、実績評価モデルは、ロジック・モデル及び、様々な評価事例より評価モデルを構築した森田モデルより導出したものであり、今後、さらなる精緻化が望まれる。

最後に、本稿では、実績評価モデルが独立行政法人制度における評価の構造を明らかにした。その結果、インプットについて、評価書に記載がないものがあることが明らかとなった。さらに、実績評価モデルの要素を確認する過程の中で、評価基準の3つの要素のうち、尺度を記述しているにも係わらず、基準値や比較対象を設定していないケースが多いという課題を検出した。このことは実績評価モデルを活用することで、評価担当者が、評価を過不足なく実施できることを示している。

実績評価モデルの活用により、実際の評価の現場において、評価者のスキルに依存することなく、一定の品質を保った評価結果を得ること

に貢献することが期待される。なお、実績評価モデルの活用によって評価基準に関する課題が把握されたが、この改善提案については、今後の課題とする。

参考文献

- 岡本信一（2001）「独立行政法人制度概説」、『季刊行政管理研究』、(96)：59-82
- 岡本義朗（2008）『独立行政法人の制度設計と理論』、中央大学出版部
- 国際観光振興機構「国土交通省所管独立行政法人の平成27年度における業務実績評価の結果について」、国土交通省、<http://www.mlit.go.jp/common/001141890.pdf>（閲覧日：2018年10月16日）
- 国立公文書館「独立行政法人の評価結果等年度評価（対象平成27年度）」、<http://www8.cao.go.jp/hyouka/doppou/hyouka.html>（閲覧日：2018年10月16日）
- 政策評価・独立行政法人評価委員会「独立行政法人総覧（平成25年度版）：国際観光振興機構」、総務省、http://www.soumu.go.jp/main_content/000315985.pdf（閲覧日：2018年10月16日）
- 政策評価・独立行政法人評価委員会「独立行政法人総覧（平成25年度版）：国立公文書館」、総務省、http://www.soumu.go.jp/main_content/000315896.pdf（閲覧日：2018年10月16日）
- 政策評価・独立行政法人評価委員会「独立行政法人総覧（平成25年度版）：防災科学技術研究所」、総務省、http://www.soumu.go.jp/main_content/000315914.pdf（閲覧日：2018年10月16日）
- 政策評価・独立行政法人評価委員会「独立行政法人評価年報（平成25年度版）」、総務省、http://www.soumu.go.jp/main_sosiki/hyouka/dokuritu_n/dokuhou_nenpou_25.html（閲覧日：2018年10月16日）
- 総務省「政策評価制度について」、http://www.soumu.go.jp/main_sosiki/hyouka/seisaku_n/000065209.html（閲覧日：2018年10月16日）
- 総務省「独立行政法人一覧（平成30年4月1日現在）」、http://www.soumu.go.jp/main_content/000544212.pdf（閲覧日：2018年10月16日）
- 総務省「独立行政法人制度等」、http://www.Soumu.go.jp/main_sosiki/gyoukan/kanri/satei2_01.html（閲覧日：2018年10月16日）
- 多賀谷一照（1998）「独立行政法人論と行政制度」、『季刊行政管理研究』、(82)：3-13
- 防災科学技術研究所「国立研究開発法人防災科学技術研究所の平成27年度における業務の実績に関する評価」、http://www.bosai.go.jp/introduction/pdf/27monka_hyoka.pdf（閲覧日：2018年10月16日）
- 森田弥生（2015）「独立行政法人制度における評価」、『日本評価研究』、15（1）：57-67
- Ackoff, Russell L. (1962). *Scientific Method : Optimizing Applied Research Decisions*. New York : John Wiley & Sons.
- Innovation Network, Inc. (2010). *Logic Model Workbook*, http://www.pointk.org/client_docs/File/logic_model_workbook.pdf（閲覧日：2018年10月16日）
- Lave, Charles A., and March, James G. (1975). *An Introduction to Models in the Social Sciences*. Lanham, Md.: University Press of America.（佐藤嘉倫・大澤定順・都築一治訳（1991）『社会科学のためのモデル入門』、ハーベスト社）
- Morita, Yayoi (2015). A Study of Inclusive Model for Creating Appropriate Project Evaluation Items, *Proceedings of the 9th International Conference on Project Management*, 387-392.
- Wholey, Joseph S. (1979). *Evaluation : Promise and Performance*. Washington, D.C.: Urban Institute.
- W.K.Kellogg Foundation (1998). *Evaluation Handbook*. <https://cyc.brandeis.edu/pdfs/reports/EvaluationHandbook.pdf>（閲覧日：2018年10月16日）
- W.K.Kellogg Foundation (2004). *Logic Model Development Guide*. <https://www.bttop.org/sites/default/files/public/W.K.%20Kellogg%20LogicModel.pdf>（閲覧日：2018年10月16日）

(2019.1.11 受理)

Construction of a Performance-Evaluation Model for Incorporated Administrative Agency System

Yayoi Morita

Graduate School of Business Sciences, University of Tsukuba
morita@gssm.otsuka.tsukuba.ac.jp

Abstract

A new Independent Administrative Agency (IAA) system appeared as a result of the reform to the Act on General Rules for IAAs. To ‘maximize policy implementation’ in the IAA system, unified government guidelines for evaluating the work performed and the goals set were presented. However, in reality, it is difficult to properly evaluate for the content of the goal description is unclear and abstract. Moreover, Evaluators in charge of the evaluation are performed within a limited time frame and a limited number of persons. As a result, evaluators feel overburdened and experience “evaluation fatigue.” However, in order to guarantee the quality of evaluation results, it is indispensable to properly incorporate the elements necessary for evaluation without overs or shorts. Based on the philosophy and reality of such an IAA evaluation, in this paper, I propose the Performance-Evaluation Model aiming to enhance Performance-Evaluation in the IAA system. By utilizing this model, evaluators will be able to detect elements that are missing in the evaluation and contribute to appropriate evaluation implementation.

Keywords

Meta-evaluation of IAAs, Morita Model, Quality Assurance of Evaluation, Component

日本評価学会第19回全国大会
「SDGsの国内展開と評価」"Localization of SDGs and Evaluation"
開催の報告とお礼

2018年12月1日、2日開催（於：横浜国立大学）の日本評価学会第19回全国大会には、152名の方々にご出席いただきました。誠にありがとうございました。会員各位の日頃の研究や実践活動の報告をもとに、評価研究者、実務者の間の経験、情報、知識を共有化する場として、皆様にとって有益な機会となりましたら幸甚に存じます。多くの方のご尽力により本大会を開催することができましたこと、心より感謝申し上げます。今後とも当学会の活動にご高配賜りますようお願い申し上げます。

実行委員長 氏川恵次／加藤郁夫（横浜国立大学）
プログラム委員長 牟田博光（東京工業大学）

2018年12月1日（土）

09：30 - 10：00	受 付		
午前の部 10：00 - 12：00	<共通論題1> SDGs評価：未来の世代に、 素晴らしい世界を2 （佐々木亮） 経済103	<自由論題1> 評価手法 （齊藤貴浩） 経済203	<自由論題2> 行政評価 （南島和久） 経済202
お昼休み 12：00 - 13：15	お昼休み (12：05 - 13：00 理事会 経済100)		
午後の部Ⅰ 13：15 - 15：15	<共通論題2> 評価の国際的潮流 （湊直信） 経済103	<共通論題3> 日本のエビデンスに基づく 政策（EBPM）への取組と 期待：EBPM導入現場の実情 （正木朋也） 経済203	<自由論題3> 自治体評価 （小島卓弥） 経済202
午後の部Ⅱ 15：30 - 17：30	<シンポジウム> 国内のSDGsに関する活動事例とその評価 （氏川恵次） 経済111		
17：40 - 18：20	総会 経済111		
18：30 - 20：00	懇親会 食堂		

2018年12月2日（日）

9：30 - 10：00	受 付		
午前の部 10：00 - 12：00	<p style="text-align: center;">＜共通論題4＞</p> <p>国際協力機構（JICA）における 新たな取り組み</p> <p style="text-align: center;">（青柳恵太郎） 経済103</p>	<p style="text-align: center;">＜共通論題5＞</p> <p>政策評価制度とその断片化 —行政学からの議論—</p> <p style="text-align: center;">（山谷清志） 経済203</p>	<p style="text-align: center;">＜共通論題6＞</p> <p>評価的思考： 日本のソーシャル・セクターに 活用重視の評価を広める鍵概念</p> <p style="text-align: center;">（今田克司） 経済202</p>
お昼休み 12：00 - 13：15	お昼休み (12：05 - 13：00 編集委員会 経済100)		
午後の部Ⅰ 13：15 - 15：15	<p style="text-align: center;">＜共通論題7＞</p> <p>国の施策・事業の評価 ～行政評価局調査の 新たな展開～</p> <p style="text-align: center;">（長澤真吾） 経済103</p>	<p style="text-align: center;">＜共通論題8＞</p> <p>国会の政策分析・評価機能 強化の為の独立財政機関 （IFI）の国会への付置</p> <p style="text-align: center;">（廣野良吉） 経済203</p>	<p style="text-align: center;">＜ラウンドテーブル1＞</p> <p>評価ディベート： 評価のお悩み相談会2018</p> <p style="text-align: center;">（小島卓弥） 経済100</p>
午後の部Ⅱ 15：30 - 17：30	<p style="text-align: center;">＜共通論題9＞</p> <p>SDGs達成に向けた評価の キャパシティ・ビルディング： グローバル目標と現場の実践を つなぐのは誰か？</p> <p style="text-align: center;">（米原あき） 経済103</p>	<p style="text-align: center;">＜ラウンドテーブル2＞</p> <p>再考「評価の研究」</p> <p style="text-align: center;">（西出順郎） 経済100</p>	<p style="text-align: center;">＜自由論題4＞</p> <p>人材育成・学校評価</p> <p style="text-align: center;">（佐藤由利子） 経済202</p>

（ ）は座長です。

日本評価学会第19回全国大会
「SDGsの国内展開と評価」 “Localization of SDGs and Evaluation”
プログラム詳細

2018年12月1日(土) 受付				9:30-10:00	
2018年12月1日(土) 午前の部 10:00-12:00					
共通論題1 SDGs評価: 未来の世代に、素晴らしい世界を2					経済103
			座長 佐々木亮	国際開発センター	
K1-1	民間企業のSDGsレポートの課題と方向性			三井久明	国際開発センター
K1-2	ボランティア-国別レビュー (VNR) 報告書から読み取るSDG17 (パートナーシップ) の進捗と課題			石田洋子	広島大学
K1-3	アジア諸国におけるSDGsモニタリング・評価の進捗と残された課題			佐々木亮	国際開発センター
自由論題1 評価手法					経済203
			座長 齊藤貴浩	大阪大学	
J1-1	EUの次期R&Iプログラム「Horizon Europe」の策定プロセスに関する一考察～我が国における次期科学技術基本計画の策定への参考素材として～			野呂高樹	未来工学研究所
J1-2	産業連関表による中国のマルチメディア産業の基礎的実証分析		○	陳泓旭	麗澤大学
				ラウシンイー	麗澤大学
				高辻秀興	麗澤大学
J1-3	体験を通じた人間関係力を育成するプログラムの評価事例		○	池田満	南山大学
				土屋耕治	南山大学
J1-4	国連気候変動枠組条約 (UNFCCC) における温室効果ガス (GHG) インベントリ作成のキャパシティ・ビルディングの効果を測る評価手法の検討		○	池田まりこ	地球環境戦略研究機関
				梅宮知佐	地球環境戦略研究機関
自由論題2 行政評価					経済202
			座長 南島和久	新潟大学	
J2-1	総務省行政評価局と行政相談の史的展開			山谷清秀	浜松学院大学
J2-2	韓国国会予算政策処 (NABO) の目的・機能・業務内容の現状と評価			上野宏	国際開発センター/神戸大学
J2-3	地方自治体におけるEBPMに基づく行政経営システムの構築とプログラム評価の方法論		○	刘谷剛	高知工科大学
				那須清吾	高知工科大学
2018年12月1日(土) お昼休み 12:00-13:15 (12:05-13:00 理事会 経済100)					
2018年12月1日(土) 午後の部I 13:15-15:15					
共通論題2 評価の国際的潮流					経済103
			座長 湊直信	国際通貨研究所	
K2-1	DAC評価項目のSDGsへの適合の議論とその背景			村岡敬一	外務省/ OECD/DAC開発評価ネットワーク
K2-2	評価に関する国際的ネットワークとアジア太平洋評価協会 (APEA) の概要			佐藤由利子	東京工業大学/アジア太平洋評価協会
K2-3	北東アジア開発協力フォーラムに参加して: 韓国、中国、ロシアのODA評価の現状と課題			石田洋子	広島大学
共通論題3 日本のエビデンスに基づく政策 (EBPM) への取組と期待: EBPM導入現場の実情					経済203
			座長 正木朋也	国際協力機構/北里大学	
K3-1	エビデンスに基づく政策立案-その承諾と本質-			田辺智子	国立国会図書館
K3-2	広島県EBPM推進ワーキンググループの取り組み			石田直人	広島県
K3-3	日本におけるエビデンスに基づく政策形成 (EBPM) の現状と課題			小林庸平	三菱UFJリサーチ&コンサルティング
K3-4	国の行政実務の立場から-EBPM推進の現状・課題と未来像			砂山裕	総務省
自由論題3 自治体評価					経済202
			座長 小島卓弥	NTTデータ経営研究所	
J3-1	参加する市民・参加しない市民2-参加型評価のために無作為抽出された市民を対象とするアンケートの結果から-自治体評価に対するイメージの役割による共通点と相違点			窪田好男	京都府立大学
J3-2	階層分析法を用いた事務事業評価の枠組み			池田葉月	京都府立大学大学院 (博士後期課程)
J3-3	介護予防活動に関連したプログラム評価開発における参加型評価活用の可能性		○	飯田洋市	公立諏訪東京理科大学
				下園美保子	愛知県立大学
				小西めぐみ	愛知県立大学
				大山早紀子	川崎医療福祉大学
				大島巖	日本社会事業大学
				松田哲子	奈良県下市町地域包括支援センター
				岸恵美子	東邦大学
				浜崎優子	佛教大学

2018年12月1日(土) 午後の部II 15:30-17:30

シンポジウム 国内のSDGsに関する活動事例とその評価		座長 氏川恵次 横浜国立大学 討論者 佐分利応貴 笹川平和財団 金藤正直 法政大学		経済111
S-1	横浜国立大学理事・副学長よりご挨拶		中村文彦	横浜国立大学
S-2	趣旨説明および論点提起		加藤郁夫	横浜国立大学
S-3	なぜ今バランス・スコアカードなのか		吉川武男	横浜国立大学/エジンバラ大学
S-4	富士ゼロックスの復興推進室の活動について		樋口邦史	富士ゼロックス
S-5	南足柄みらい創りカレッジとSDGsに関する活動について (南足柄みらい創りカレッジより中継) (南足柄みらい創りカレッジより中継) (南足柄みらい創りカレッジより中継)		遠藤一弥	富士ゼロックス
			鈴木正	富士ゼロックス
			富田直子	ウィルウィンド
	パネルディスカッション		黒柳俊之	国際農業開発

2018年12月1日(土) 総会 17:40-18:20 (経済111)

2018年12月1日(土) 懇親会 18:30-20:00 (食堂)

2018年12月2日(日) 受付 09:30-10:00

2018年12月2日(日) 午前の部 10:00-12:00

共通論題4 国際協力機構(JICA)における新たな取り組み		座長 青柳恵太郎 メトリクスワークコンサルタンツ		経済103
K4-1	JICAの事後評価における最新動向		新井和久	国際協力機構
K4-2	事後評価の定量分析:説明変数の定義と実務課題	○	南和江	国際協力機構
			小川一弥	グローバルリンクマネジメント
			正木朋也	国際協力機構
K4-3	JICAとフィリピン政府との合同評価による一考察		田中優子	国際協力機構
K4-4	JICAのプロセスの分析～事業の運営・管理者による評価の活用に向けて～	○	坂井美保子	国際協力機構
			平田桃	慶応義塾大学大学院(修士課程)

共通論題5 政策評価制度とその断片化—行政学からの議論—		座長 山谷清志 同志社大学 司会/討論者 西山慶司 山口大学		経済203
K5-1	評価制度の断片化によるアカウントビリティの散逸		山谷清志	同志社大学
K5-2	大学の組織と評価—そのディメンジョン—		南島和久	新潟大学
K5-3	行政と評価関係—研究開発評価・政策評価・独立行政法人評価—		橋本圭多	神戸学院大学

共通論題6 評価的思考:日本のソーシャル・セクターに活用重視の評価を広める鍵概念		座長 今田克司 CSOネットワーク 司会 源由理子 明治大学		経済202
K6-1	なぜ評価的思考か、そしてそれはいかに実践されるのか—今日の状況における重要性(英語/逐次通訳)		Archibald, Thomas	Virginia Tech University
K6-2	日本のNPO/ソーシャル・セクターで、なぜ今ETに注目することが必要か		今田克司	CSOネットワーク
K6-3	発展的評価とET		中谷美南子	評価コンサルタント

2018年12月2日(日) お昼休み 12:00-13:15 (12:05-13:00 編集委員会 経済100)

2018年12月2日(日) 午後の部I 13:15-15:15

共通論題7 国の施策・事業の評価～行政評価局調査の新たな展開～		座長 長澤真吾 総務省 討論者 山谷清秀 浜松学院大学 司会 橋本圭多 神戸学院大学		経済103
K7-1	最近の行政評価局調査の事例—いじめ防止対策の推進に関する調査—		長谷川章	総務省
K7-2	最近の行政評価局調査の事例—買物弱者対策に関する実態調査—		中澤京介	総務省
K7-3	最近の地域計画調査の事例—障害のある学生等に対する大学の支援に関する調査—		山根京子	総務省

共通論題8 国会の政策分析・評価機能強化の為の独立財政機関 (IFI) の国会への 付置				座長 廣野良吉	成蹊大学	経済203
K8-1	日本版独立財政機関の提言案(叩き台)と米国事例		上野宏	神戸大学/国際開発センター		
K8-2	IFI導入上の我が国における課題		湊直信	国際大学		
ラウンドテーブル1 評価ディベート: 評価のお悩み相談会2018 ~研究・実践活動 において日々生じている評価の課題や疑問 (あるいはお悩み) を共有し、解決の糸口を探る~				座長 小島卓弥	NTTデータ経営研究所	経済100
				進行補助 西出順郎	明治大学	
2018年12月2日(日) 午後の部 II 15:30-17:30						
共通論題9 SDGs達成に向けた評価のキャパシティ・ビルディング: グローバル目標と現場の実践をつなぐのは誰か?				座長 米原あき	東洋大学	経済103
				司会 源由理子	明治大学	
K9-1	SDG Evaluation: Translating Global Level Concepts to Local Level Actions		RUGG, Deborah H.	Claremont Graduate University		
K9-2	Practices and problems of localizing SDGs-Goal 4.7: Case studies at school context		米原あき	東洋大学		
K9-3	How to build ownership by people is at the heart of localizing the SDGs		今田克司	CSOネットワーク		
ラウンドテーブル2 再考「評価の研究」~「評価の研究」のあり方について議論を 深める~				座長 西出順郎	明治大学	経済100
				討論者 大島巖	日本社会事業大学	
				山崎その	京都外国語大学	
自由論題4 人材育成・学校評価				座長 佐藤由利子	東京工業大学	経済202
J4-1	ミャンマー軍事政権下の初等教育における CAPSの効果検証		吉田夏帆	関西学院大学大学院(博士課程)		
J4-2	「2015年洪水及び地滑り被害地における学校復旧計画」実施が児童生徒数の変動に与えた影響の評価分析		牟田博光	国際開発センター/大妻女子大学		
J4-3	コミュニティスクールにおける学校評価ニーズについての調査-中部・関東地方の公立学校の場合-		橋本昭彦	国立教育政策研究所		
J4-4	JICA「青年研修」事後評価と研修効果の要因分析	○	岸本昌子	日本国際協力センター		
				菊田玲子		

日本評価学会第19回全国大会セッション報告

シンポジウム

シンポジウム 「国内のSDGsに関する活動事例とその評価」

企画 加藤 郁夫（横浜国立大学）
 座長 氏川 恵次（横浜国立大学）
 討論者 佐分利 応貴（笹川平和財団）
 金藤 正直（法政大学）

はじめに中村文彦氏（横浜国立大学）より、横浜国立大学で日本評価学会の初めての大会を開催することに歓迎のお言葉を頂いた。また同学は、近年地域貢献の面において国内で評価が高まっていることを述べられ、さらに地域連携の一環としてのこのような産学連携のシンポジウムの意義をご示唆頂いた。続いて加藤郁夫会員（横浜国立大学）より、本シンポジウムの趣旨として、横浜国立大学と包括連携協定を締結している富士ゼロックス株式会社及び神奈川県南足柄市により開校された、一般社団法人南足柄みらい創りカレッジを事例として取り上げ、その活動のアウトカムを評価する方法として、バランス・スコアカード（以下BSC）を取り上げるという内容の説明を行った。

第一講演は吉川武男氏（横浜国立大学）より、「なぜ今バランス・スコアカードなのか」というタイトルで、SDGsのようなグローバルな目標への取組を考慮せざるを得ない現代社会において、BSCを利用することで、社会に貢献するアウトカムを導くことができるという内容のご講演を頂いた。

第二講演は、樋口邦史氏（富士ゼロックス株式会社、一般社団法人南足柄みらい創りカレッジ）より「南足柄みらい創りカレッジの挑戦－SDGsを掲げた地域活性化の実践－」というタイトルで、広域防災とSDGsを二大テーマに、協働的な実践活動をまち創りの指針としてプログラムを設計しているという内容のご講演を頂いた。企業がそのような活動を行う背景には、グローバルレベルの社会環境と、ビジネスを取り巻く関係に配慮した経営が求められているからであり、組織のTOPは自ら社会的インパクト評価を実践し、経営戦略に役立てる必要があると述べられた。BSCに関しては、そのような手法をグループ単位で実施することも重要であるとのことご講演を頂いた。ご講演途中、南足柄みらい創りカレッジと中継で結び、鈴木正氏（富士ゼロックス株式会社復興推進室、南足柄未来づくりカレッジ）及び富田直子氏（有限会社ウイルウインド代表取締役）から会場の様子をレポートして頂いた。なお、カレッジの目指すSDGsの姿は、黒柳俊之氏（株式会社国際農業開発）が描いた「私たちの夢～2020」をイメージしていることから、黒柳氏にご出演頂いた。

第三講演は、遠藤一弥氏（富士ゼロックス株式会社、南足柄未来づくりカレッジ）より、事例紹介として、南足柄みらい創りカレッジ「みんなで学ぼうSDGsプログラム」に関するBSCについてご講演を頂いた。「魅力あるSDGs実践リーダー育成プログラムを開発し、SDGs実践リーダーを育成し、SDGs達成に向けた様々な実証実験を行える最先端の実験フィールドを整備する」というビジョンを示された。

パネルディスカッションでは、佐分利応貴会員（公益財団法人笹川平和財団）から、SDGsの国内での各取組を、単発で終わるのではなく普及可能なモデルに発展させ横展開し、ソーシャル・イノベーション（＝解決策：solution × 普及：diffusion）にすることが必要であるとのことご意見を頂いた。また、取組のインパクトを評価するには具体的な指標やツールが必要であり、BSCの有効性を示す事例蓄積を期待したいとのコメントを頂いた。次に、金藤正直会員（法政大学）より、『地域事業へのSDGsとBSCの導入ポイント』

と題してお話を頂いた。日本には、サプライチェーン（またはバリューチェーン）や産業クラスターなどのように、当該事業のコアとなる人材や組織（コーディネーターなど）が、当該地域の自治体、民間事業者、地域住民・NPOなどといった事業関連組織を連携させ、地域再生を目指した取組を持続的に行う地域事業を展開している事例は決して多くない。SDGsを導入する場合でも、持続可能な地域事業を行い、事業成果を獲得し続けていくためには、先述のコーディネーターが事業関係組織とともに、CSV（Creating Shared Value）の概念やSDG Compassを加味しながら、地域再生のための経営システムの構築、すなわち、経済価値と環境・社会的価値をともに創出する地域事業のための戦略を策定し、その戦略を実行させるための最適な組織編成や管理プロセスの構築が必要であると述べられた。BSCは、地域事業でのこうした経営システムを有効的かつ効率的に運営・管理していくためのモニタリング・システムとして機能させることが重要であるとのコメントを頂いた。

共通論題セッション

共通論題1 「SDGs評価：未来の世代に、素晴らしい世界を2」

座長 佐々木 亮（国際開発センター）

2016年から始まった「持続可能な開発目標」（SDGs）は、貧困に終止符を打ち、地球を保護し、すべての人が平和と豊かさを享受できるようにすることを目指す普遍的な行動を呼びかけている。本セッションでは、三つのトピックについて報告して論じ、今後のSDGsの実現について展望した。

冒頭、座長より趣旨説明及び共通論題セッションの趣旨と論点の整理を行った。

続いて、「研究報告1：民間企業のSDGsレポーティングの課題と方向性」として、国際開発センターの三井久明氏から報告があった。続いて、「研究報告2：ボランティア国別レビュー（VNR）報告書から読み取るSDG17（パートナーシップ）の進捗と課題」と題して、広島大学の石田洋子氏から報告があった。さらに、「研究報告3：Progress and Remaining Issues of Monitoring and Evaluation on SDG among Selected Asian Countries（アジア諸国におけるSDGsモニタリング・評価の進捗と残された課題）」として、国際開発センターの佐々木亮氏から報告があった。

それぞれの報告のあとに、クラリフィケーションの質疑応答があり、三つの発表のあとにフロアからコメント及び質問をいただいた。それぞれの発表に対して有意義なコメント及び質問があった。特に、「報告者の所属する機関である国際開発センターでは、SDGsへの取り組みをどのようにレポーティングしているか」という質問があり、現在、Global Reporting Initiative（GRI）という専門機関からSDGsレポーティングに関する講師の認証を受けた職員を複数名誕生させて、今まさにレポーティングに取り組んでいるという説明があった。

共通論題2 「評価の国際的潮流 (International Trends of Evaluation)」

座長 湊 直信 (国際通貨研究所)

本国際セッション「評価の国際的潮流」では海外で国際開発分野の評価に関して行われている様々な議論を共有することを目的にした。現在、国際機関、援助機関、政府、大学、学会、NGOs、評価専門家ネットワーク等では、様々な形で評価について議論が行われている。同時に開発途上国の評価能力の現状や、評価制度や評価システムの進展も我々の大きな関心事である。最近、海外の評価学会や評価専門家の会合に参加した発表者から、評価に関する最新の議論や新しい考え方、アプローチ、及び評価能力の進展等について報告が行われた。それに基づいて、質疑応答が行われた。

第一報告は村岡敬一会員 (外務省) による「DAC評価項目のSDGsへの適合の議論とその背景」である。SDGs時代の到来により、長年、開発協力評価の基準となっていたDAC評価5項目の有用性に関する議論が展開されている。現在、OECD/DACでは2030アジェンダに適合する評価項目のありかたの検討を進めており、2019年春までに結論を得ることになっている。本報告においては、DAC評価項目を巡る国際的な議論の背景とDAC開発評価ネットワーク (EvalNet) の取組と進捗状況が紹介された。DAC評価項目ワークショップで論じられたDAC評価5項目の強みや課題の紹介、評価項目に関する今後の方向性が述べられた。

第二報告は佐藤由利子会員 (東京工業大学) による「評価に関する国際的ネットワークとアジア太平洋評価協会 (APEA) の概要」である。現在、国単位の評価学会の他に団体をつなぐ地域単位の評価ネットワーク組織、更にこれらを含む世界的なネットワーク組織が形成されている。評価に関する数多くの国際的なネットワーク組織の概要の説明と、2012年に設立されたアジア太平洋評価協会 (APEA) の組織概要と主な活動の現状が報告された。2019年2月に予定されている第2回APEAコンフェレンスに多くの日本評価学会員が参加するよう、情報共有が行われた。

第三報告は石田洋子会員 (広島大学) による「北東アジア開発協力フォーラムに参加して: 韓国、中国、ロシアのODA評価の現状と課題」である。第5回「北東アジア開発協力フォーラム」は、「SDG達成へ向けての北東アジア諸国間の評価における国際協力」をテーマとして2018年9月にソウルで開催された。日頃、日本ではあまり情報のない韓国、中国、ロシアの評価制度に関する現状や課題等が報告された。日本のODAの評価経験への具体的な関心等についても述べられた。同時に南と南の国際協力に焦点を当てて、南のニーズを反映させた国際協力のあり方についての意見も紹介された。

全ての報告終了後に質疑応答が行われた。主な論点は、援助側と被援助側の評価への関心、中国の評価制度に対する動機、APEAの途上国人材の評価に関する経験と背景等々であった。

共通論題3 「日本のエビデンスに基づく政策 (EBPM) への取組と期待: EBPM導入現場の実情」

座長 正木 朋也 (国際協力機構/北里大学)

日本のEBPMが多様な解釈のもとで推進されつつある現状について、先の春季第15回全国大会 (2018/05/26、東京) で取り上げ共有を図ってきた。このセッションでは改めて国内のEBPM展開にあたり、その推進に直接携わる各現場担当者を招聘し議論を行った。

冒頭で座長は趣旨説明に加えて、EBPM展開の歴史、エビデンスにはレベルと質があること、証拠に基づく医療 (EBM) の意思決定モデル、因果推論における定量・定性的アプローチのバランスなどエビデンススペースの理解に関わる基本事項を共有した。

第一報告者の田辺智子会員（国立国会図書館）は「エビデンスに基づく政策立案—その系譜と本質—」と題し、EBPMを展開するうえで必要不可欠な基本事項について概念整理を行った。EBPMの原型とされるエビデンスに基づく医療（EBM）を参照しつつ基本事項を説明し、因果推論との関係からエビデンスを理解することが肝要であること、政策形成におけるエビデンスが立場に応じて多様な意味と解釈のもとで用いられ、その混同による無用の議論が生じ得る状況があることを指摘した。

第二報告者の石田直人氏（広島県）は「広島県EBPM推進ワーキンググループの取り組み」と題し、広島県の事例として、重要性・緊急性の高い領域のエビデンスを作るための仕組みと組織的な取り組みを紹介した。現場のモチベーション向上の好循環を起こすことの重要性に触れ、国や自治体でEBPMを推進する際、こうした現場レベルでの活動事例をもとに相互に意見交換する機運が醸成されてゆくことへの期待も述べた。

第三報告者の小林庸平氏（三菱UFJリサーチ&コンサルティング）は「日本におけるEBPMの現状と課題」と題し、EBPM導入推進者らを支援する立場から俯瞰的に現状を分析した。統計情報の利活用推進から始まった日本独自のEBPMの特徴を捉え、そのような状況で精緻な因果推論を前提としたエビデンスの議論を行うことの難しさがあることを紹介した。その状況を踏まえた課題解決にあたり、エビデンスを「つたえる」「つかう」ことよりも現時点では「つくる」ための仕掛けづくりが優先されるとし、また、その需要の創出を促す政府側の手腕も必要になるであろうと考察した。

第四報告者の砂山裕氏（総務省、賛助会員）は「国の行政実務の立場から—EBPM推進の現状・課題と未来像」と題し各省庁のEBPMを推進する立場からの報告を行った。政府は、方法論やツールからではなく、需要やどのようなことがより良い政策立案につながるかといった本質的な問いから考えることが肝要とした。その際、人材開発、研究機関やコンサルタントの活用、データの利活用、及び倫理規範への配慮などの条件整備を進めることも必要であり、戦略的な判断とともにハイレベルの構想も必要であるとまとめた。

参加者からは、政策評価や支援事業の中で「ついでにエビデンスを作る」といったことは実務上不可能であるので、どこで誰がどのように作るべきかの意見交換がなされた。また、どのようなエビデンスが必要とされるかに関わる合意形成の方法やそのエビデンス作成のための新たな枠組みが必要であること、さらにそれらを伝えるためのデータベース等のインフラ整備の必要性について、誰がどのように予算や人材を確保可能であるかといった運用に関わる観点からの議論もなされた。また、EBPMを議論する際その評価に繋がる示唆が必要との意見もあったが、科学的なエビデンスの話と政策形成でそれらを活用し評価する話は本来別途議論すべきものではないかとの意見もあった。

以上、この度のセッションは冒頭と第一報告の概念整理も功を奏し、異なる立場の参加者間においてもすれ違いなく活発な議論が行われ、会場の定員を超える参加者らによる熱い議論はセッション終了後もフロアにて継続されたことを付記しておきたい。

共通論題4 「国際協力機構（JICA）における新たな取り組み」

座長 青柳 恵太郎（メトリクスワークコンサルタンツ）

国際協力機構（JICA）が行う政府開発援助は体系化された評価システムを有しており、国外からも高い評価を受けている。そうしたJICAの評価について、これまでの全国大会でも最新の取り組み、課題等について共有を図ってきた。そこでは単なる事例紹介にとどまらず、実践現場から提起される論点の評価論への位置づけ、解釈を行ってきた。今次全国大会のJICAセッションでは、評価論の発展過程を念頭に、「現在の国際開発領域の課題に対して、評価がどのように共鳴しあいながら変容を遂げきているのか」という

視点から各報告で扱われる事例を取り扱っていくことを試みた。

第一報告は、「JICAの事後評価における最新動向」と題して、新井和久氏（JICA評価部）より国内の評価関係者を念頭にJICA評価部の考え方・姿勢が伝えられた。この後に続く個別の取り組み事例の土台となるJICAの評価全般を俯瞰する報告となった。様々な論点に触れる報告となったが、なかでもJICA事業評価における評価の中立性、評価実施プロセスにおける評価者への関係者の反応といった点で、評価倫理との接点を見出された。ここでは評価倫理を担保する、ないしは評価倫理が蹂躪されてしまう事態を回避するための仕掛け・仕組み、心構えといった議論に展開を見せた。また、評価が果たせる学習（ラーニング）とは何か、換言すれば、評価に期待してはいけない学習とは何か曖昧になってきており、評価の守備範囲を過度に拡張していないかという問題提起があった。

第二報告は、南和江氏（JICA評価部）、小川一弥氏（グローバル・リンク・マネジメント）、正木朋也会員（JICA評価部）より「事後評価の定量分析：説明変数の定義と実務課題」として、JICA評価部が取り組む個別事業のレーティングに関する統計分析について、分析に用いる変数作成プロセスを通じて直面した課題や解決策についての紹介があった。個別の事業評価にレーティングシステムを導入している開発援助ドナーは多くはないが、その代表的機関である世界銀行は古くからレーティングの決定要因を探る研究を行ってきた。JICAの取り組みもこうした研究の流れの中に位置づけることができる。その分析に供する変数の定義と作成・選択や加工、モデル選択などで分析者がどのような実務上の課題を抱えているのかは通常語られることがない。本報告は、分析の根幹となるデータ構築の実情が共有される貴重なものとなった。

続いて、田中優子氏（JICA評価部）より、スタンドバイ借款事業でフィリピン政府と合同評価を行った事例とそこからの考察を扱った「JICAとフィリピン政府との合同評価による一考察」報告があった。ドナー国とレシピエント国が共同して行う合同評価は開発援助特有の評価形態である。本報告に対しては、座長より、これまで合同評価は、①先方政府の評価能力強化、②学びの機会が協働により深まること、③説明責任の断絶（broken information feedback loop）が克服できることが論じられてきたとの整理がなされた。そして、本報告はこれらを踏まえた合同評価の事例紹介という点に加えて、プログラムローン・一般財政支援の評価をどう行うべきかという評価手法の議論としても意義があったことが強調された。

最後に、「JICAのプロセスの分析～事業の運営・管理者による評価の活用に向けて～」と題して、坂井美保子氏（JICA評価部）、平田桃氏（慶応義塾大学）よりJICAが新たに実施する「プロセスの分析」について、その概観と他組織が実施する類似の手法についての調査結果が報告された。プロセス「の」分析が一般的な評価論（プログラム評価）との関係でどのように位置づけられるのか、評価としての要素をどう持ち合わせているのかといった点で、非常に野心的な取り組み事例報告であった。

国際開発が取り組む発展途上国の課題、実施環境は日々目まぐるしい変化を示しており、評価システムにも様々な工夫や新たなアプローチが取り入れられていく。実践者の取り組みの中に、新たな評価論の展開が見いだされることは疑いない。こうした取り組みを断片的な事例として扱うのではなく、評価論の中に位置づけた議論を今後も追い求めていく必要がある。

共通論題5 「政策評価制度とその断片化—行政学からの議論—」

座長 山谷 清志（同志社大学）

司会／討論者 西山 慶司（山口大学）

これまで日本の政府では政策評価をはじめとして数々の評価制度が登場してきたが、それらは十分に整理、統合されているのではない。政策評価を制度化する過程で生じてきた「評価の断片化」状況とどのように向かい合うべきなのか。この問いを考えるため、本セッションでは、「政策評価の制度化」「個別政策

領域での断片化」「断片化した評価間関係」という三つの観点から議論した。

第一報告は山谷清志「評価制度の断片化によるアカウンタビリティの散逸」である。評価の目的の一つはアカウンタビリティの追及であるが、日本では政策を担当する行政機関が自らアカウンタビリティを確保するように制度化された。その結果、行政機関はみな「説明すれば責任を果たしたことになる」と理解した。他方で、同じ行政機関は評価をマネジメントに使う。いわゆるマネジメント型評価で、効率や有効性、コンプライアンスがPDCAの中で重視される。しかもここに、もう一つの評価の目的、専門分野に対する知的貢献が入る。専門評価である。科学や医療など高度に専門的な分野は一般市民に理解が難しく、それでもマネジメント評価する難しい仕事が行政機関職員自身に発生する。目的と趣旨が違う評価関係業務が重畳的に重なり、評価疲れという病理現象がこうして生じてくる。

この問題提起に関わる具体事例として、南島和久会員（新潟大学）は「大学の組織と評価－そのディメンション－」と題して、「個別政策領域での断片化」の実例を提示する。①大学進学者の大幅な減少への対応、②学位プログラムを拡充したい大学・学部の希望と教育組織再編に関わる設置認可等の縛りの矛盾、③学部・研究科の3ポリシー（アドミッション、カリキュラム、ディプロマ）の自己点検自己評価との①②との不整合、④学部・研究科3ポリシー評価と学位プログラム評価、組織評価との分裂、これらの四つの問題が大学評価に生じて、日本の高等教育政策の評価は断片化しているのである。

類似の問題は研究開発評価にも発生する。その具体的事例は第三報告の橋本主多会員（神戸学院大学）の「行政と評価間関係－研究開発評価・政策評価・独立行政法人評価－」である。この議論の前提は政策評価と研究開発評価、独立行政法人評価（研究開発法人評価）の併置である。それだけでも整理が難しく評価業務の難解さが精神的負担になるが、さらに行政事業レビューの対象になった時、研究開発の当事者（科学者と技術者）は、何をどのようにすれば成果を出したと証明でき、良い評価を得られるのか悩ましい。結局、アカウンタビリティとして何を、いつ、どのように説明できれば合格点が得られるのか、それが収斂すべき「落としどころ」を評価を命じる評価主体、評価実務担当者、評価される客体、評価を見る側（市民とその代表）がイメージしていないため不信感が出てくる。

以上の報告に対して、西山慶司会員から政策評価、独立行政法人評価、大学評価の不整合に関するコメントがあり、評価自体もセクショナリズム、「タコ壺」化に陥っているのではないかと質問を受けた。それぞれ報告者がこの質問に答えた後、フロアの他の会員ともディスカッションがあったが、最終的には「評価は民主主義のリテラシーである」の原則に立ち返り、一般市民に理解できる評価システムを再構築する必要を確認してこのセッションは終わった。

共通論題6 「評価的思考：日本のソーシャル・セクターに活用重視の評価を広める鍵概念」

座長 今田 克司（CSOネットワーク）
司会 源 由理子（明治大学）

本セッションでは、評価的思考（Evaluative Thinking）に焦点を当て、米国バージニア州のメイン報告者とオンラインでつなぎ、ビデオ会議方式で報告をしてもらった。逐次通訳は座長が行った。約40人が参加し、熱の込められた議論の場となった。

メイン報告者のトマス・アーチボルド氏（バージニア・テック大学）は、北米において評価による「学び」を促進する鍵概念として（再）浮上している評価的思考について概説した。全米評価学会の学会誌 *New Directions for Evaluation*（2018年夏号）は、評価的思考の特集号を組んでいる。評価専門家のあいだでは「当たり前」としてあまり意識されない評価的思考だが、現代的課題に応える重要な思考のありようと

して注目すべきものとなっている。アーチボルド氏は、なぜこれが大事なのか、いかにこれを意識・強化できるかを解説した。

「なぜ？」という問いに対しては、1) 社会のエビデンス志向が強まるにつれ、評価の役割に対する関心が高まっており、2) 「価値づけ」、「価値の引き出し役」としての評価が注目されており、3) 評価キャパシティビルディングの動き、協働型・参加型評価の実践の増大などの評価の民主化が進んでおり、4) 発展的評価に見られるような複雑系理論の評価への応用が進展しているからという四つのポイントが示された。これらはすべて評価という営みに内在されている評価的思考の役割や重要性を浮かび上がらせているという。

一方の「いかに？」では、一例として、評価の対象である事業の運営や運営組織、関係者がもっている暗黙の想定・了解事項を掘り出すことや、事業に関わる関係者それぞれ（そこに権力関係が存在する）がもつ特定の立場、見方を可視化することが紹介された。

評価的思考を養い、有効活用するには、仕事のルーチンなど日常レベルでこれを意識する組織文化づくりが大事だということが強調された。また、評価的思考を評価者以外が持つことも重要である一方、評価活動においても評価的思考が効いていないこともあるとの論評があった。

続けて行われた質疑応答では、社会福祉評価や学校評価の専門家などから、評価的思考が必要とされている分野が多様に広がっているという認識が共有された。

ビデオ参加のアーチボルド氏が退出したあと、日本で評価的思考を意識した評価活動がいかに可能かを話し合った。解題&コメント提供者として、中谷美南子氏（評価コンサルタント）から、「評価的思考のキャパシティディベロップメントとしての発展的評価」の発表があった。CSOネットワークによる発展的評価の事例づくりの一環で行ったNPOでの価値基準の可視化、セオリーオブチェンジの特定などを目指したワークショップが、評価的思考醸成のための「足場かけ」と位置づけられた例が紹介された。

最後に座長より、現在のNPO／ソーシャル・セクターでの評価の位置づけと評価的思考の将来的役割が紹介され、今後の議論の広がりに対する期待感が共有された。

共通論題7 「国の施策・事業の評価 ～行政評価局調査の新たな展開～」

座長	長澤 真吾（総務省）
討論者	山谷 清秀（浜松学院大学）
司会	橋本 圭多（神戸学院大学）

総務省行政評価局は、国の事業・施策について、「政策評価」、「行政評価局調査」及び「行政相談」の三つの機能を発揮して行政上の課題の解決を図り、もって国民に信頼される質の高い行政を実現することを目指している。本セッションで取り上げたのは、これらのうち「行政評価局調査」である。

「行政評価局調査」については、府省の業務の実施状況を实地に調査し、各府省の課題や問題点を実証的に把握・分析し、改善方策を提示するものであり、全国的な調査（全国計画調査）と地域の行政上の課題についての調査（地域計画調査）がある。

第一報告者（長谷川章氏 総務省行政評価局）からは、1) 「最近の行政評価局調査の事例－いじめ防止対策の推進に関する調査－」として、いじめ防止対策について、教育現場の参考となるよう学校等が工夫している取組事例の紹介を中心とするという、従来の行政評価局調査とは異なるアプローチにより実施された調査の特徴に言及しつつ、地方公共団体による自殺等の重大事態に関する調査報告書の分析や教育現場の実態を把握した同調査について報告がなされた。

第二報告者（中澤京介氏 総務省大臣官房）からは、2) 「最近の行政評価局調査の事例－買物弱者対策

に関する実態調査」として、買物弱者対策という、所管府省がなく、対策の全体像も明確でないテーマについて、改善方策の提示を行うことにこだわらずに、その全体像と課題を明らかにすることを主眼に置くという、従来の行政評価局調査とは異なるアプローチにより実施された調査の特徴に言及しつつ、国や地方公共団体における買物弱者対策に資する事業の実施状況等を把握した同調査について報告がなされた。

第三報告者（山根京子氏 総務省近畿管区行政評価局）からは、3)「最近の地域計画調査の事例－障害のある学生等に対する大学の支援に関する調査－」として、①外部有識者の専門的知見の活用、②障害のある方々・大学とのバリアフリー共同調査、③視覚障害者等によるHP点検、④障害のある学生に対する意識調査という新たな取り組みに言及しつつ、国立大学法人等の支援に関する実態等を把握した同調査について報告がなされた。

コメンテーターの山谷清秀会員からは、報告のあったそれぞれの行政評価局調査について、①これらの調査が実施された背景や過程、②これらの調査が従来とは異なる点の詳細に関する質問がなされたほか、1) に関し、行政評価局調査において政策目的と現場のギャップを指摘することの意義等について、2) に関し、①改善方策の提示（勧告）を行わないとした決定のプロセス、②調査により把握した情報を共有するという事例が行政評価局調査においてどの程度見られるものであるのか等について、3) に関し、①改善意見の効力、②改善意見の提示後のフォローアップの方法、③当該調査の他の管区への波及効果等について質問がなされ、これらに対する応答がなされた。併せて、フロアからは、行政評価局調査のテーマ選定過程における政策評価審議会での議論の状況等に関する質問がなされた。

共通論題8 「国会の政策分析・評価機能強化の為の独立財政機関（IFI）の国会付置提案とその課題」

座長 廣野 良吉（成蹊大学）

日本の累積債務は、本年GDPの約2.4倍に達し、OECD諸国で最大である。来年度の一般会計当初予算案は戦後初めて100兆円を超え、今後更に増大すると予想されている。主な原因は、経済成長優先のための低金利・ゼロ金利政策と連動した赤字財政政策と特に各種既得権益集団への助成・補償という名目での放漫な財政支出による。この巨大な借金、我々の子・孫・ひ孫が将来返済せねばならず、現世代による将来世代の搾取を意味する。国債費は年々増大し、持続可能な財政、特に国民生活の安定・改善に必要な投資支出はおろか、消費税引き上げなくしては、経常支出の維持さえ困難な状況に追い込まれている。少子高齢化の進展により、将来世代の負担は今後益々増大することが予測される。

この問題の解決には、行政府が国会へ提出する予算案に内包される財政支出・歳入政策を厳しく分析・評価する必要がある。多くのOECD諸国では既にその対策として国会が法制化を通じて赤字財政規模の上限を設けたり、厳しい予算審議を通じて財政規律の再生・強化に努めたりしている。一つの手段が独立財政機関（IFI）の設立であり、米国では、既に議会予算局（CBO）が長年機能しており、EUはその加盟国にGDPの3%という制限を予算に義務づけている。日本もIFIないしその類似機関の設置が喫緊の課題となっている。IFIは同時に、国民の直接選挙により選出された国会議員の関与で、行政府の行動の監視・監理を通じて、三権分立を基本とする民主主義政治体制の強化にも役立っている。本年度の通常国会、臨時国会でも、衆参両議院で予算委員会が開催されたが、本格的な予算内容についての議論がほとんど見られなかった。その背後には、政権与党議員はそれぞれが所属する党内政策調査会・分科会などでの事前審議を通じて、政府原案の策定過程に深く関与しているが、同じく有権者によって選出された野党議員の大半は、行政府が年々国会へ提出する予算案が、政府が掲げる優先政策ないし重要政策との整合性、合理性につい

て検討する「時間と場」を十分に持っていないことに起因しているといっても過言ではない。かかる現状に鑑み、政党とは無関係であらゆる予算関連情報を入手可能で、分析能力・意欲を持って国会予算委員会における審議を補佐するIFIの有効性は高いと考えられる。さらに、IFIの設立により、予算案の背景にある経済・社会政策に関する独立的、科学的、客観的分析により、選挙法改定により「小選挙区の利害」のみ偏りがちの与野党議員が、国民経済全体への関心度を高めることに繋がり、国民にとっての意義も大きい。究極的には財政規律の徹底、日本経済の持続的・安定的な成長、国民各層の経済・社会福祉の改善への貢献を最終主眼とするIFIが、国会議員の指向・選好の転換に向けてどの程度有効か、さらに地元の利害に敏感かつ関心が深い有権者、住民の内向きの考え方を、より広く国民全体の利害への関心を高めるという「パラダイムシフト」にとって、どの程度有効かということの検証も不可欠である。

日本評価学会は、我が国に「評価の文化」を根付かせて、当事者内外による政策・施策・事業評価の普及・改善へ貢献することを主目的の一つとしている。政策分析・評価を目的とするIFI設置をマクロ、ミクロの経済・社会政策や制度の分析・改善を主目的とする各種学会との協力・連携の下で推進する最適学会である。そこで本学会では、2018年3月「立法府の分析評価機能強化に関する分科会」を設立した。分科会は4月から勉強会を重ねているが、「評価の文化」が未だしという我が国、特に政治家の間では、本提案は未だ優先課題となっていないのが現状である。

司会者より、本セッション企画の背景、趣旨の説明に引き続いて、報告者2名により、主要な論点について簡潔な報告が行われた。その後司会者によるいくつかの質問と付加的な説明に続いて、参加会員との間での質疑応答・討論を行った。今回のセッションでは、予想以上に多くの会員が参加しただけでなく、若手の女性会員が多かったのが印象的であった。

第一報告者、上野宏会員（国際開発センター）は、主に①我が国がIFI導入を必要とする理由を、上記導入の背景に結び付けて説明し、特に財政破綻が引き起こす経済社会の混乱を回避するための財政の持続可能性の達成を重視した。次に②IFIの目的に言及し、特に三権分立の中での最高決定機関である国会による財政に関する分析・評価機能を強化する必要性を説き、中長期に及ぶ財政の将来を推計し、世代間の財政負担の公正化と縦割り行政による個別政策予算の重複・無駄の削減を図ると共に、分析・評価情報を広く国民一般へ公開し、国民の参加を促し、民主主義制度を守り強化することの重要性を訴えた。さらに、③IFIの業務内容（機能）と組織に言及して、政党・政府・行政・その他全ての外的影響を受けずに、内閣から提出される予算案とそれが支える政策案を国会の予算委員会による審議以前に、できうる限りの客観的・科学的な分析・評価を行い、その結果を予算委員会をはじめ各種委員会へ提供し、国会がこの情報を基により客観的で厳密な予算審議を可能にすることの重要性を訴えた。また、かかる情報提供のためには、IFIは、必要な場合には国会又はその委員会に要請し、政府や民間機関に対して国会の国政調査権による情報収集を要請できると共に、機密情報の守秘義務は当然である。なお、効率的かつ効果的な分析・評価のために、120名程度の専門家を擁し、マクロ経済・歳出部、歳入・税政策部、と3種の部門別プログラム分析部を持つことを提案した。IFIと会計検査院との職務・機能の重複問題については、会計検査院が予算執行後の事後評価を行うのに対して、IFIは予算審議前に予算案の事前評価を行うこと、政策評価法による政策評価は個別政策の事後評価の内部評価であり、IFIはむしろ予算全体の評価で、事前評価・外部評価であること、衆・参議院の行政監視委員会は、メンバーが国会議員で、所属党派や議員個人の方針に従った評価であるのに対して、IFIは、専門家による独立で客観的な分析・評価機関であることが強調された。最後に、IFI導入の期待される成果については、特に国会の予算案・政策案の分析・評価能力の強化への貢献を挙げ、国会議員への情報と全く同じ情報を広く国民・メディアへ公表することで、国民の予算・財政への関心、政治参加を高め民主主義制度の改善・強化に寄与すること、行政府、特に内閣・財務省の予算案形成に緊張感を与えることができることを強調した。また、IFIの普遍性、実効性を強調する意味で、OECD諸国、特に米国議会に所属する議会予算局（CBO）の目的、機能、組織について説明し、大統領府に設けられた予算管理局（OMB）との違い、国会に予算提出権がある米国と議院内閣制を採用している我が国

との違いを指摘して、米国流のCBOに近い形でのIFIの優位性を説いた。

第二報告者の湊直信会員（国際大学）は、「IFI導入のわが国における課題」をテーマに、近年のGDP、財政赤字、失業率等のデータと共に、現時点でIFIを日本に導入する際に予想される課題と解決方法を説明した。主要な論点として、①財政赤字に対する国民各層間における認識の違い、②国会議員一般の問題意識の欠如、③行政府、特に財務省の予算編成権に対する介入懸念、④IFIの内容が抽象的で理解しにくい、⑤他国の情報を含めて、IFIに関する情報が不十分、⑥IFIの有用性について議論する機会が少ない、⑦議会制民主主義体制維持に不可欠な三権分立下での立法府による行政府活動へのチェック機能の重要性に対する認識不足、⑧IFI導入への活動主体の欠如、⑨赤字財政支出の上限設定が年金や社会保障制度の崩壊につながるという懸念、という9点に絞った簡潔な説明は説得力があった。若干詳述すると、①日本の財政状況についての専門家等の意見の焦点が経済成長率、失業率、インフレ率等それぞれ違うために、財政赤字に対する認識は共通ではない点を重視し、今後は日本の財政状況の深刻さについて認識を共有することの必要性を説いた。②2013年に9名の衆参両議員が、将来の議員立法を念頭に「国会に独立将来推計機関の設置」を提案したが、民主党から自民・公明両党への政権交代もあり、多くの与野党国会議員は本提案に積極的に関わらなかった。この状況は今日も続いているため、国民、メディア、オピニオンリーダー等への広範な啓蒙活動の必要性を力説した。③議院内閣制の日本では国家予算は行政府（主に財務省）で作成され、国会に提示される。IFI導入には行政府、特に財務省と官邸が反対する可能性もあるが故に特に財務省との意見交換を通じて、財政の中長期的な持続性の重要性を共有することが必要。④OECDが示しているIFIに関する原則案（オーナーシップ、独立性と非党派性、権限（役割）、財政基盤、立法府との関係、政府情報へのアクセス、透明性、評価等）を今後日本で実現するための具体的な行動計画（ロードマップ）の作成が必要だという指摘があった。⑤他の先進国の事例を紹介することで、具体性、実現性のあるIFIを示すことができる。さらに、住民に直結した地方自治体の大半が赤字財政を抱えている中で、地方自治体におけるIFI導入の成功例を提示できれば、国レベルでの導入に拍車をかけることが想定される。⑥財政赤字の問題だけでなく、憲法改正、消費税の引き上げ等その他国民の関心が大きい課題についての議論の機会に、IFIについての議論を展開することが望ましい。⑦民主主義体制の維持発展には、その受益者である国民各層の意識改革を含めた不断の努力が必要であり、国会による国会付置のIFI導入に関する討論はこの重要な課題について格好の機会を提供すると共に、日本に「評価文化」を根付かせる啓蒙活動の一環でもある。⑧IFI設立のための各種活動を牽引していく独立した組織の設立を早急に検討することが望ましい。⑨IFIの将来推計が年金等の社会保障給付の状況も含むことにより、国の債務と年金給付の推計シミュレーションによる選択肢を提示し、国会議員をはじめ、国民各層の間での議論を促すことを強調した。日本評価学会の会員が経済学や財政金融の専門家と共にIFI導入に向けて、しっかりしたロードマップを作成することが最も近道であるという結論は説得力があり、参加した多くの会員が賛同していた。

2名の報告に引き続いて、参加者との討論の中心は、IFIの具体的な中身と共に、国会に付置された各種委員会や現行の会計検査院との関係、特に会計検査院の権限強化によって代替できないのかという質問もあった。IFIそのものについての議論もさることながら、IFI導入に横たわる諸々の実践的課題や、従来から導入に消極的ないし否定的な与党、政治家個人々人、行政府官僚への説得を如何に進め、IFI導入の最大の難関を如何に乗り越えるかが最大の課題であるという認識が報告者と参加者の共通の最大関心事であったとあって良いであろう。この点は分科会の座長として勉強会でも観察してきた議論と相通ずるところであり、WinWinのIFI導入戦略の模索を加速化し、勉強会における徹底的な議論を通じて、実行可能なロードマップ作成の責任を強く感じた。その意味からも、今回の全国大会セッションでは、先に分科会の勉強会でその必要性を議論した。我が国の数百人に及ぶ各界の有権者への財政赤字とIFIに関する意見を求めるアンケート調査案のサンプルを参加者へ配布し、年内の回答を依頼し、今後国会議員をはじめ一層多くの日本評価学会会員の分科会、勉強会への参加を訴えた。本セッション企画者としては、単に、「IFI提言案（叩き台）」の中身を精査するだけでなく、その実現を可能にする具体的なロードマップ作成に向けて、特

に若手の会員による活発な意見交換を期待したが、時間的制約もあり、その重要性和原則論、方向については合意したものの、IFI導入の具体的なロードマップの大枠の作成という目標を達成できなかったことは残念であった。また、一部の会員については、本件への関心が未だという印象も拭いえない。しかし、わが国の将来設計におけるIFIの重要性に鑑み、分科会による勉強会を今後も続けると共に、国会議員や行政、学者等を主対象とした活動のみならず、国会議員を選出する各地域の有権者個人、集団への教宣活動が如何に重要かを再認識し、主権者教育の全国展開の重要性を改めて深く認識した。

共通論題9 “Evaluation capacity building for SDGs: Who connect the global goals with local practices?” (SDGs達成に向けた評価のキャパシティ・ビルディング：グローバル目標と現場の実践をつなぐのは誰か?)

座長 米原 あき (東洋大学)
 司会 源 由理子 (明治大学)
 討論者 Rugg, Deborah H. (Claremont
 Evaluation Center - New York,
 Claremont Graduate University)
 米原 あき (東洋大学)
 今田 克司 (CSOネットワーク)

従来のグローバル目標に比して、SDGsには一層抽象度の高い目標が含まれていることが広く知られている。これらの抽象的な概念を現実に実施し、効果的に評価を行うためには、評価者と現場の関係者が協働して、これらのグローバル目標をローカルな文脈へ翻訳する必要がある。このようなニーズにこたえるため、評価のキャパシティ・ビルディングが喫緊の課題となっている。本セッションは、グローバル及びローカルな観点から、評価のキャパシティ・ビルディングについて検討することを目的として企画された。本セッションでは、元・国連評価部上級アドバイザーのデボラ・ラグ教授にメイン・スピーカーとしてご登壇頂き、その後、SDGsのローカル化と評価的思考 (evaluative thinking) の観点から二つの報告を行い、フロア・ディスカッションの機会を設けた。

デボラ・ラグ教授 (クレアモント大学院大学) による基調講演では、“*Evaluation Capacity Building: How to Translate Global Goals into Local Action*” と題する講演が行われた。MDGsからSDGsへより複雑化した国際目標の達成を実現するためには、それらの複雑な目標を評価することのできる人材の育成・確保と制度の適切な設計が不可欠となる。特に近年、政府や省庁のみならず、市民団体や公共セクター、大学やメディアなど、評価に関わるステイクホルダーは多様化している。これらの潮流を踏まえ、ラグ教授は、コロンビアにおける評価のキャパシティ・ビルディング案件を紹介した。コロンビアではSINERGIAと呼ばれる、国家開発計画のマネジメント制度を導入し (Stage 1: 1994-2002)、ここに評価の機能を持たせた。続くStage 2 (2002-2014) においてモニタリング評価のガイドラインを策定するなど、モニタリング評価の強化と制度化を進め、Stage 3 (2014-2030) では、SDGsの評価を実施するべく内閣にハイレベル委員会を設置している。この事例及びその他の関連事例から、キャパシティ・ビルディングには時間がかかるということ、国ごとの文脈がキャパシティ・ビルディングと評価デザインにとって非常に重要な要素であること、各関係者との協働や情報共有が肝要であること、その文脈のなかで既存のリソースや制度を活用したキャパシティ・ビルディングを行う必要があること、そして、あらゆるレベルにおいてステイクホルダーからのインプットが重要であることが明らかになった。同時に、「指標は必要だがそれだけでは充分ではない」という指摘がなされ、「評価=指標」と捉えられがちな傾向に対して警鐘が鳴らされた。

続いて米原あき会員（東洋大学）は、“*Practices and Problems of Localizing SDGs-Goal4.7: Case Studies at School Context*”と題する報告のなかで、SDGsの中でも評価が困難であると考えられている目標4.7に注目し、特に抽象度が高く曖昧な概念であるESD（Education for Sustainable Development：持続可能な開発のための教育）の評価について議論した。ESDのような理念的な目標は、そもそもそれが具体的に何を意味するのかを、各文脈の実生活のレベルで再定義する必要がある。ユネスコがunpackingと呼ぶこの再定義のプロセスを、参加型のセオリー評価として実施することで、学校現場の関係者たちがこの再定義のプロセスに関わることができ、現場の関係者のアイデアから指標が生まれるなど、学校関係者にとって有用なプロジェクト活動及び評価活動が実現することが、いくつかの事例から示された。

最後に今田克司会員（CSOネットワーク）が、“*How to Build Ownership by People is at the Heart of Localizing the SDGs*”と題された報告において、SDGs達成のための取り組みに関しては、地域のオーナーシップの醸成が鍵となることを論じた。今田会員は、「3つのステップ」—即ち、①日本の地方地域（local community）はSDGsに従事する準備ができていることを認識する、②SDGsの取り組みに関して政府と地方地域の連携を推進する、そして③SDGsに取り組む意義の理解を促進し、地方地域がより体制を整えるために、地域の人々を巻き込む—を示したうえで、現在日本に存在する27のSDGs未来都市の可能性に言及した。特に「3つのステップ」を実施するうえで重要なのは、これらを実践のキャパシティ・ビルディングの活動として行うことであり、この過程で人々の評価的思考が培われ、また、異なるステイクホルダーたちが目的意識を共有するようになることが期待される。このような評価活動は昨今注目を集めているdevelopmental evaluationの考え方と親和性が高いことも指摘された。

本セッションでは、研究者のみならず、国際協力機関や自治体関係者などの参加も得て、活発な議論が交わされた。研究のうえでも実践のうえでも課題は山積しているが、抽象的なグローバルゴールをローカルレベルに落とし込む際に、参加型評価や評価的思考が鍵となり得ることが確認された。これらの点に関する今後の研究・実践の展開に期待したい。

ラウンドテーブル1 「評価ディベート：評価のお悩み相談会2018」

座長 小島 卓弥（NTTデータ経営研究所）
進行補助 西出 順郎（明治大学）

本セッションは、第18回全国大会に引き続いて開催され、評価の理論と実務の有機的結合を具現化する取り組みの一つとして実施したものであり、2回目の取り組み（試み）となる。

学会員各位の日々の研究もしくは実践活動において日々生じている課題や疑問（あるいはお悩み）を広く共有し、また、当該議論を多様な活動領域や実務、理論といった複眼的なレンズを通じ、これらの解決の糸口を模索するものである。

前回同様、特定の評価領域を設けず、政策評価、自治体評価、NPO評価、施設評価、EBPM等幅広い分野における評価活動における課題、疑問、問題意識が寄せられ、それぞれに対し、参加者から自らの研究、あるいは実践経験から助言が供されるなど、評価の現場で発生している様々な課題に対して活発な議論が行われた。

評価の実践者は相談する相手がおらず、一人で苦悩するケースも少なくないようだ。このような場を通じてそれぞれが抱える課題に対し、研究者と実践者がそれぞれの知見を持ち寄り、課題解決を図ることは非常に重要であると考えている。

また、本学会において分野、研究者、実践者の壁を越えて議論が行われることは、我が国における評価の発展という側面からも非常に重要だと考えており、このような機会を継続することができればと考えている。

ラウンドテーブル2 「再考「評価の研究 (The study of Evaluation)」」

座長 西出 順郎 (明治大学)
 討論者 大島 巖 (日本社会事業大学)
 山崎 その (京都外国語大学)

本RTDは、「評価」を研究対象として考察するにあたって、その分析視角や手法等のありかたを改めて議論し、その全体像を探ることを目的としたものである。

はじめに大島巖会員 (日本社会事業大学) 及び山崎その会員 (京都外国語大学) から話題提供が行われ、その後フロアを交えて議論が展開された。

評価にかかる一連の用語の定義、記述的アプローチと実態分析アプローチ等の研究客体への接近法、応用科学としての評価実践への貢献、評価活動報告書論文の位置付け、研究結果と外的妥当性の関係等、多様な論点から、各論者それぞれの学問領域に依拠した評価の研究としてのディスプリンを共有できたと考える。

勿論、全体像を描くという意味では、論点を拡散させた故に議論としては十分に尽くされたわけではなく、ここで明示することすらままならぬが、評価を研究対象とする上での最大公約的ベースラインを可視化する意義はフロア間では共有できたと考える。

自由論題セッション

自由論題1 「評価手法」

座長 齊藤 貴浩 (大阪大学)

本セッションでは四つの研究報告に対し、フロアからの質疑を得て活発な議論が行われた。なお、発表者は、当日登壇した方のみを記載する。

まず、第一報告者である野呂高樹会員 (未来工学研究所) の「EUの次期R&Iプログラム「Horizon Europe」の策定プロセスに関する一考察～我が国における次期科学技術基本計画の策定への参考素材として～」では、EUが2021年から2027年までの7年間で展開する研究&イノベーションプログラムの策定プロセスが詳細に報告された。産学官の各種利害関係者からの意見聴取に基づき、現行プログラムからの教訓や、将来予測、経済的裏付けなどの検討を行い、社会的インパクトをもたらす道筋とそのモニタリング指標を設け、オープンな環境での実施が謳われており、日本の科学技術基本計画の策定に比べ、計画段階の事前評価の手厚さや意思決定の方法の違いが大きく、日本側の政策形成の遅れが指摘された。

第二報告者である陳泓旭会員 (麗澤大学) の「産業連関表による中国のマルチメディア産業の基礎的実証分析」では、中国の産業連関表から、日本の基本分類に沿ってマルチメディア産業として28の部門を抽出し、その特性についての分析を試みた報告がなされた。マルチメディア産業では企業側と雇用者側の所得格差が殆どないこと、中国の市場経済の特徴が認められること、最終需要に影響を与えるのは金融業や教育業であること、高等教育による人材需給は十分であることなどが指摘された。本学会では産業連関表

からのアプローチはあまり行われておらず、経済活動の把握のためには数値の意義をより詳細に見ていく必要はあるが、さまざまな活用が期待される。

第三報告者である池田満会員（南山大学）の発表は、「体験を通じた人間関係力を育成するプログラムの評価事例」であり、大学の教育プログラムの評価の事例研究であった。EIAHE' モデル（体験、指摘、分析、仮説化、（再）体験）という体験学習のモデル、及び尺度を援用し、質問紙調査の結果を分析したところ、体験学習の事前事後で社会的スキルと仮説化の得点の上昇に有意差が認められた。本研究は長期間にわたり展開してきた講義であり、多くの知見とデータが累積できていることから、質疑応答の中でも複数年度の分析を通じての効果の違いや、考え方の違いなどへのアプローチの有効性についてのコメントがあった。

最後に、第四報告者として池田まりこ会員（地球環境戦略研究機関）より、「国連気候変動枠組条約（UNFCCC）における温室効果ガス（GHG）インベントリ作成のキャパシティ・ビルディングの効果を測る評価手法の検討」と題した報告があった。温室効果ガスインベントリ（GHGI）とはUNFCCCの締約国がそれぞれに示すべきGHGの実績であり、国際機関やドナーの協力で作成され、同時に各国のキャパシティ・ビルディング事業が行われており、その事業評価を行おうとするものであった。現在は主なドナーのキャパビル事業の特徴を示すに留まり、詳細な評価はこれからとのことであり、事業評価として有意義な結果がもたらされることが期待される。

本セッションでは、「評価手法」を切り口に、様々な分野の発表が行われた。評価手法は必ずしも一つの分野に固有の手法があるわけではなく、様々な適用可能性がある。今後、分野を超えて横断的に評価手法や問題の解決策を探ることも必要になるだろう。今後の進展と深化が期待される。

自由論題2 「行政評価」

座長 南島 和久（新潟大学）

本セッションは自由論題として三つの報告を扱い、あわせて会員相互の討論を行った。なお、セッションの標題である「行政評価」は、「行政部内の評価」というような広い意味であることを補足しておきたい。

第一報告は山谷清秀会員（浜松学院大学）による、「総務省行政評価局と行政相談の史的展開」であった。その内容は、2017年10月に行われた総務省の地方支分部局の再編に関するものであった（全国の「行政評価事務所」の一部を「行政監視行政相談センター」に変更し、各事務所が有していた行政評価の機能が各管区行政評価局に集約するもの）。同機構改革にあわせて「地域総括評価官」のポストが管区及び各センターに新設され、各管区レベルでの行政評価の機動性の向上が図られている。山谷報告は、歴史的経緯も踏まえつつこの新たな動向を紹介し、論評するものであった。総務省の新しい動向を共有する点に本報告の意義を見いだすことができるだろう。

第二報告は、上野宏会員（国際開発センター／神戸大学）による、「韓国国会予算政策処（NABO）の目的・機能・業務内容の現状と評価」であった。上野会員はこれまで、IFI（独立財政機関）の研究を継続的に本学会にて発表してきており、本報告はその一環として捉えられるものである。上野報告によれば、NABOは世界第二位の規模のIFIであり、2012-2016年度国会審議では予算減額の7割をNABOの意見を反映するなど良好なパフォーマンスを発揮しているということであった。上野報告の結論は、国民一般への啓蒙が不足している点及び目的・機能と主要業務との不整合等を指摘するものであった。上野会員は、他の会員とも協力して日本版IFIの設立についての社会的活動を展開しているところであり、その基礎となる議論を蓄積するという意味において本報告は重要な意義があるといえるだろう。

第三報告は、刈谷剛・那須清吾会員（高知工科大学）による、「地方自治体におけるEBPMに基づく行政

経営システムの構築とプログラム評価の方法論」であった。本研究は、「エビデンス・ベースによる戦略的行政経営システムの構築」の開発と自治体への実装を目指すものである。事例については、自治体で行われている特定健康診査の実施率の分析・評価手法が取り上げられていた。会場においては、ロジックモデルの前提となるプログラムの明確化やエビデンス・ベースとデータドリブンの相違などが話題となった。公的部門においてはEBPMに話題が集中しているところ、本研究のような個別事例に注目した評価手法に関する議論は重要性を増していくものと見込まれる。

自由論題3 「自治体評価」

座長 小島 卓弥 (NTTデータ経営研究所)

本セッションでは、「自治体評価」と題して、地方自治体における、行政評価、及びそれ以外の評価に関して、四つの研究及び実践の報告がなされた。

第一報告者の窪田好男会員（京都府立大学）からは「参加する市民・参加しない市民2－参加型評価のために無作為抽出された市民を対象とするアンケートの結果から－」と題して、舞鶴市において参加型評価を実施するために無作為抽出された候補（参加・不参加双方を含む）に対して、アンケート調査を実施し、参加・不参加の要因、どのような属性の方が参加されるのか、また参加しやすくする方法等について調査・分析結果が報告された。質疑応答では、市民参加型評価により職員にとりストレスが軽減されたというが、その結果の評価の効果は本当にあるのか、という質問や市民参加型評価の根拠や意義等、市民参加型評価そのものに対する質問が寄せられた。

第二報告者の池田葉月会員（京都府立大学）からは「自治体評価に対するイメージの役職による共通点と相違点」と題して、宝塚市の室長（課長と部長の間の職位）に対して、行政評価の実施イメージについて漢字一文字で表現してもらい、かつ評価に対する印象を回答してもらうアンケートを実施し、過去に実施した課長・係長クラスの回答傾向との差等について調査・分析結果が報告された。質疑応答では、役職の差ではなく、部署の差があったのか（宝塚市の例では見られなかった）、サンプル数が少ないのではないかと、他自治体でも同様の調査を実施し、同じような傾向がみられるか確認してはどうか、等の質問が寄せられた。

第三報告者の飯田洋市会員（公立諏訪東京理科大学）からは「階層分析法を用いた事務事業評価の枠組み」と題して、茅野市における男女共同参画事業について、階層分析法に基づき階層図を設定し、かつそれに対して個別の評価値等を設定し、階層分析法に基づく分析の試みについて報告があった。質疑応答では、階層分析法の実施方法や、そもそもの階層構造の組み方や見せ方等、階層分析法の実施方法等について質問が寄せられた。

第四報告者の下園美保子会員（愛知県立大学）他からは「介護予防活動に関連したプログラム評価開発における参加型評価活用の可能性」と題して、地域介護領域における介護予防活動におけるプログラム理論の一部開発を目的に、住民を含む各ステークホルダーが参加した参加型評価を実施した実践例と、その成果に対する有効性や効果について報告された。質疑応答では、ロジックモデルの構成や、アウトカムの設定理由、製作したロジックモデルに対して、設定しにくい成果指標をどのように把握していくのか、等の質問が寄せられた。

いずれの発表にも多くの質問が寄せられ、参加者の関心の高さが感じられた。

地方自治体における行政評価はせっかく導入されたにも関わらず、総務省の調べでは既に5%程度が廃止もしくは休止したという。一方で、今回の四つの発表のように、様々な形で行政評価あるいは、行政評価の枠の外にある評価の取り組みは継続的に実践・研究されている。今後も本学会において、これらの研究と実践の交流が盛んに行われ、より実効性の高い自治体評価へと発展・昇華させていくことが期待される。

自由論題4 「人材育成・学校評価」

座長 佐藤 由利子（東京工業大学）

本セッションでは、ミャンマーにおける初等教育の学力評価制度の効果検証、災害被害地における学校復旧計画の評価、コミュニティスクールにおける学校評価ニーズについての調査結果、JICA青年研修の効果の要因分析など四つの貴重な報告があり、活発な質疑応答が行われた。

吉田夏帆会員（関西学院大学）による「ミャンマー軍事政権下の初等教育における CAPSの効果検証」では、1998年に子どもたちの学力を継続的に評価する制度として導入されたContinuous Assessment and Progression System（CAPS）による修学状況の改善への影響を、対象校所在地域の「社会経済的地位SES」別に、時系列に分析し、SESが高位～中位の地域では修学状況に改善が見られるものの、SESが低位の地域では十分な改善が見られないことが報告された。質疑応答では、学校のタイプ（小中高一貫校か小学校のみの学校か）、対象地域における民族や宗教的特徴などについて質問があった。

牟田博光会員（国際開発センター／大妻女子大学）による「[2015年洪水及び地滑り被害地における学校復旧計画] 実施が児童生徒数の変動に与えた影響の評価分析」では、2015年のミャンマーにおける洪水及び地滑り被害地における学校復旧計画支援の効果について、児童生徒数の増減を指標として、傾向スコアマッチング手法により評価した結果が紹介された。質疑応答では、JICAが実施した財政支援方式の利点と課題、評価結果から導かれる政策的含意などについて質問があった。

橋本昭彦会員（国立教育政策研究所）による「コミュニティスクールにおける学校評価ニーズについての調査－中部・関東地方の公立学校の場合－」では、「学校運営協議会」を設置するコミュニティスクール（全国で3,600校）における学校評価の実態や、保護者、地域住民、学校側のニーズについて、学校評価書や学校管理職研修会でのアンケート調査に基づき分析した結果が報告された。質疑応答では、「学校関係者評価」が教育現場でどのように機能しているのか、「学校関係者評価」を実施することの多い学校運営協議会の構成員と校長との関係性などについて質問があった。

岸本昌子会員（日本国際協力センター）と菊田怜子会員による「JICA“青年研修”事後評価と研修効果の要因分析」では、青年研修の効果と、効果に影響を及ぼす要因を明らかにするため、2010年～2014年に研修を受けた590名の質問紙調査を分析した結果が示された。質疑応答では、青年研修の内容や分野、行政事業レビューによる青年研修評価との関連性などについて質問が行われた。

日本評価学会 春季第16回全国大会のご案内

春季第16回全国大会
実行委員長 正木 朋也（国際協力機構）

来たる2019年5月25日（土）に、日本評価学会春季第16回全国大会がJICA地球ひろば（東京都新宿区）にて開催されます。

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

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

記

1. 日程：2019年5月25日（土）
2. 場所：JICA 地球ひろば
〒162-0845 東京都新宿区市谷本村町10-5

以上

日本評価学会誌刊行規定

2005.2.15改訂

2002.9.18改訂

2001.9.9改訂

(目的および名称)

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

(編集委員会)

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

(編集方針)

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

- (5)「その他」には、編集委員会が独自に企画する特集に掲載する依頼原稿や学会誌の刊行に関する編集委員会からの学会会員への連絡等が含まれる。
13. 論文等は2名の査読者により査読することとし、その人選は編集委員会が行う。「研究論文」については、査読結果と編集委員会が査読者とは別に指名する担当編集委員1名の参考意見をもち、編集委員会が掲載に関する決定を行う。「総説」、「研究ノート」、「実践・調査報告」および「その他」の論文については、査読結果にもとづき編集委員会が掲載に関する決定を行う。
14. 編集委員が「評価研究」に投稿した場合には、当該委員はその投稿に係わる常任編集委員会あるいは編集委員会の議事に一切参加しないものとする。
15. 上記いずれのカテゴリーの投稿についても、常任編集委員会による掲載の判断は可・不可の二者択一で行うこととする。但し、場合によっては編集委員会の判断で、小規模の修正による掲載も認める。「研究論文」としての掲載が適当でないと判断された場合でも、投稿者が希望すれば、常任編集委員会は「研究ノート」あるいは「実践・調査報告」としての掲載を決定できる。
- (投稿要領の作成公表)
16. 編集委員会は、上記の編集方針にもとづき投稿要領を作成し、理事会の承認を得て、広く公表する。
(配布先)
17. 「評価研究」は、学会会員に無償で配布するほか、非会員に有償で提供する。
(抜刷の配付)
18. 「評価研究」掲載論文等の抜刷り30部を、投稿者（原著者）に無料で配布する。それ以上の部数を希望する場合は投稿者（原著者）の自己負担とする。
(インターネット上の公開)
19. 「評価研究」掲載論文等は、投稿者（原著者）の了承を得て全文をインターネット上で公開する。
(著作権)
20. 「評価研究」に掲載された論文等の著作権は各投稿者（原著者）に帰属するものとし、編集権は本学会に帰属するものとする。
(事務局)
21. 「評価研究」編集及び配布の事務は、それに関連する会計も含めて学会事務局が担当する。

(以上)

『日本評価研究』投稿規定

2019.1.21改訂

2017.7.20改訂

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、原稿の種類が、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) 使用言語は日本語または英語とする。
 - (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) 下記送付先に投稿原稿の電子ファイル(図表を入れた印刷モードのPDFとMS word)をメールで提出する。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、最終原稿の電子ファイル(MS word)を提出する。その際、オリジナル図表(Excel等)を添付すること。なお、編集に関する最終的

なデザインは編集委員会が行うものとする。

- (5) 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。最大文字数20,000字の「研究論文」と「実践調査報告」の場合は、図表を含め刷り上がりが14ページ以内とする。ただし、15～16ページの2ページは著者負担で増頁できる。最大文字数15,000字の「総説」と「研究ノート」の場合は、図表を含め刷り上がりが12ページ以内とする。ただし、13～14ページの2ページは著者負担で増頁できる。左記最大ページ数（16ページと14ページ）を超える投稿は認めない。なお、刷り上がりとは、第1ページの和文要約、第2ページ以降の本文等（図表も含め、20字×43行×2段で設定）、最終ページの英文要約を入れた印刷モードのことである。
- (6) 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。ページ数の上限については、日本語原稿の場合と同様である。

8. 送付先

〒108-0075

東京都港区港南1-6-41

品川クリスタルスクエア12階（一財）国際開発センター内

日本評価学会事務局「日本評価研究」編集担当

E-mail: jes.info@evaluationjp.org

『日本評価研究』執筆要領

2019.1.21改訂

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（1ページ20字×43行×2段で設定する）：本文（図表含む）、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

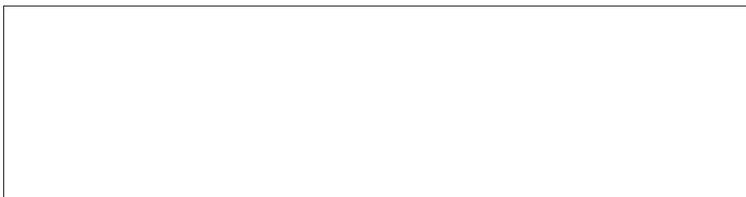
1.1.1

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

英語単行本：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、重点が異なります。

- (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
ままっていること)
- 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無い。
- 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 administers 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 administer 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 20th July 2017

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 manuscripts via email. Contact information including mailing address, telephone number, fax number, e-mail address, and the category of the manuscript should be clearly stated.
 For approved manuscripts, after necessary rewriting, the author needs to submit the final paper via email. Original figures, charts, and maps should be provided.
 - (5) Total printed pages should not exceed 14 pages. Any cost incurred by printing more than 14 pages must be covered by the author.
 - (6) The layout for English papers should be 30 mm of margin at left and right side, 10pt for font size, 43 lines on A4 paper (about 500 words per page). An abstract of 150 words should be attached to the

front. 14 pages are equivalent to 7,000 words but the body should not exceed 6,000 words to allow for the title, header, figure, chart, footnotes, and references. Please note that the number of pages may be more than expected depending on the number of figures included.

8. Mailing address

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

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

Revised on 18th September 2002

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

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

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

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

(2) Section of the text should be as follow:

1.

1.1

1.1.1

1.1.2

(3) Source of the charts, figures, graphs, and diagrams should be clarified. Submitted charts and others will be pphotoengraved, 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 Francisco, 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*. Washington: 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 Francisco, 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
〒108-0075 東京都港区港南1-6-41 品川クリスタルスクエア12階
一般財団法人国際開発センター内
E-mail : jes.info@evaluationjp.org

『日本評価研究』第19巻第2号
2019年3月29日

編集・発行 特定非営利活動法人 日本評価学会
〒108-0075 東京都港区港南1-6-41
品川クリスタルスクエア12階
一般財団法人国際開発センター内

印刷 株式会社 研恒社

©日本評価学会

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

Japanese Journal of Evaluation Studies

Vol. 19, No. 2, March 2019

CONTENTS

Special Issue: Use and Influence of Evaluations

Special Issue: Use and Influence of Evaluations Satoko Tanabe

The Use and Influence of Performance Measurement:
Empirical Analysis of Public Libraries Satoko Tanabe

Evaluation Use and Influence: A Review of Theoretical and Empirical Research
Naoko Masuda

Article

Construction of a Performance-Evaluation Model for Incorporated
Administrative Agency System Yayoi Morita

Report of the 19th Annual Conference

Announcement for the 16th Spring Conference