

日本評価研究

The Japanese Journal of Evaluation Studies

Vol. 6, No. 1, March 2006

特集：エビデンスに基づく評価の試み

特集にあたって 佐々木 亮

エビデンスに基づく医療(EBM)の系譜と方向性：
保健医療評価に果たすコクラン共同計画の役割と未来 正木 朋也 津谷 喜一郎

刑事司法におけるエビデンスの活用：
非行の予防・処遇を中心として 津富 宏

エビデンスに基づく教育
アメリカ教育省What Works Clearinghouseの動向 田辺 智子

ODA分野における『エビデンスに基づく評価』の試み：
「貧困アクションラボ」の動向 佐々木 亮

研究論文
都市大気汚染政策における社会的能力の評価 村上 一真 松岡 俊二

インドネシア前期中等教育における地方分権化の影響
教育環境と教育成果に関する因果モデルによる評価 戸井 敦子 牟田 博光

構造調整借款20年間のレビューからみる日本政府の政策と判断 田中 弥生

標準化した効果単位を用いる定量的評価手法の提案
留学生政策と地域展開型プログラムへの適用例 佐藤 由利子

情報流通に関わる政策の評価と分析
エコマーク事業を事例にして 平原 隆史

地方自治体における評価の波及と生成過程の分析 古川 俊一 森川 はるみ

研究ノート
A Study of Evaluation Methodology for Donor-funded School Construction
Projects in Developing Countries - through Establishing Evaluation Guidelines -
Satoshi Morita, Yasuyuki Sagara

実践・調査報告
学生評価インターン制度の拡充に向けた一考察
「ラオスのこども」の組織運営に関する評価調査を事例として 茂木 勇

日本評価学会第6回全国大会開催報告

日本評価学会

The Japan Evaluation Society

『日本評価研究』編集委員会

Editorial Board

編集委員長
Editor-in-chief

三好 皓一(立命館アジア太平洋大学)
Koichi MIYOSHI

副委員長
Vice-Editors-in-chief

西野 桂子(ジーエルエム・インスティテュート)
Keiko NISHINO

古川 俊一(筑波大学)
Shun'ichi FURUKAWA

常任編集委員
Standing Editors

牟田 博光(東京工業大学)
Hiromitsu MUTA

山谷 清志(同志社大学)
Kiyoshi YAMAYA

編集委員
Editors

青山 温子(名古屋大学) 大沢 真理(東京大学)
Atsuko AOYAMA Mari OSAWA

岡本 義朗(株)JFJリサーチ&コンサルティング)
Yoshiaki OKAMOTO

小野 達也(鳥取大学) 窪田 好男(神戸学院大学)
Tatsuya ONO Yoshio KUBOTA

佐々木 亮(ウェスタンミシガン大学) 佐藤真理子(筑波大学)
Ryo SASAKI Mariko SATO

渋谷 和久(国土交通省) 鈴木 絲子(立命館アジア太平洋大学)
Kazuhiisa SHIBUYA Itoko SUZUKI

田中 弥生(東京大学) 松岡 俊二(広島大学)
Yayoi TANAKA Shunji MATSUOKA

村松 安子(東京女子大学)
Yasuko MURAMATSU

事務局
Office

〒135-0047 東京都江東区富岡2-9-11 京福ビル
財団法人国際開発センター内
日本評価学会事務局
TEL: 03-3630-8031, FAX: 03-3630-8095
E-mail: jes_secrt@idcj.or.jp

日本評価研究

第6巻 第1号 2006年3月

目次

特集：エビデンスに基づく評価の試み

佐々木 亮

特集にあたって.....1

正木 朋也 津谷喜一郎

エビデンスに基づく医療（EBM）の系譜と方向性：
保健医療評価に果たすコクラン共同計画の役割と未来3

津富 宏

刑事司法におけるエビデンスの活用：
非行の予防・処遇を中心として21

田辺 智子

エビデンスに基づく教育 アメリカ教育省What Works Clearinghouseの動向31

佐々木 亮

ODA分野における『エビデンスに基づく評価』の試み：
「貧困アクションラボ」の動向43

研究論文

村上一真 松岡 俊二

都市大気汚染政策における社会的能力の評価55

戸井 敦子 牟田 博光

インドネシア前期中等教育における地方分権化の影響
教育環境と教育成果に関する因果モデルによる評価71

田中 弥生

構造調整借款20年間のレビューからみる日本政府の政策と判断85

佐藤 由利子	
標準化した効果単位を用いる定量的評価手法の提案	
- 留学生政策と地域展開型プログラムへの適用例	103
平原 隆史	
情報流通に関わる政策の評価と分析	
- エコマーク事業を事例にして	119
古川 俊一 森川 はるみ	
地方自治体における評価の波及と生成過程の分析	133
研究ノート	
Satoshi Morita, Yasuyuki Sagara	
A Study of Evaluation Methodology for Donor-funded School Construction Projects in	
Developing Countries through Establishing Evaluation Guidelines	147
実践・調査報告	
茂木 勇	
学生評価インターン制度の拡充に向けた一考察	
「ラオスのこども」の組織運営に関する評価調査を事例として	161
第6回全国大会開催報告	
第6回全国大会プログラム（実績）.....	177
共通論題セッション報告	182
自由論題セッション報告	188
委員会活動	
企画委員会	197
研修委員会	198
広報委員会	199
国際交流委員会	200
春季第3回全国大会のご案内	201
特定非営利活動法人日本評価学会定款	202
日本評価研究刊行規定	212
日本評価研究投稿規定	214
日本評価研究執筆要領	216
日本評価研究査読要領	219

【巻頭言】

特集：エビデンスに基づく評価の試み

特集にあたって

佐々木 亮

ウェスタンミシガン大学

「エビデンスに基づく評価」(Evidence-based Evaluation)が、世界の評価研究における大きな流れになっている¹。簡単に言えば、確かに効果があるというエビデンス(証拠)が確認された介入行為を政府の政策として採用してもらうことを目指す評価活動ということである²。この背景には以下のような考え方があるとされる。

「過去50年間にわたり、医学の分野においては、科学的に厳密なエビデンス(証拠)に基づいた公共政策により健康に関して大きな進展が見られた。これに比べて、例えば、教育、貧困削減、犯罪と司法、薬物乱用防止策といった社会政策の分野においては、エビデンス(証拠)はほとんど注目されずに施策が実施されていて、その結果、巨額な資金が使われているが、私たちの社会が本当に必要とすることに対して何ら解決策を見出せないでいる。しかしながら、いくつかの非常に効果がある社会的な介入(Intervention)がどのようなものであるかが、厳密な研究により明らかになってきている。このことは、これらの効果があると証明された介入の例を集め、広く利用されるような方策を政府がとれば、医学を変えたように社会政策においても急速な進展が見られるかも知れないということを意味している。」³

ここに、評価活動が社会を大きく改善するポテンシャルが垣間見える。そのポテンシャルを現実のものとするためには、上記の説明にもあるように、次の3つの作業が必要だと言える。

1. 効果があると客観的に証明された介入の評価事例の産出
2. そうした介入の事例の収集と体系化の作業
3. 広く利用されるような仕組の整備

まず「1. 効果があると客観的に証明された介入の事例の産出」であるが、これはいわゆる実験デザイン(Experimental Design)を用いるということである⁴。サンプル集団を、無作為割当(例えばコインの裏表)によって2つのグループに分け、片方のグループに介入を施し、一定期間経過後に、2つのグループに現れた違いを測定する。無作為割当によって2つのグループのあらゆる要素(例えば平均身長や男女比率)が等しくなっているわけで、2つのグループに現れた差は、途中の唯一の違いである介入を実施したかしなかったかの違いによって引き起こされたと考えることができる。つまり2つのグループに現れた差は、純粋に介入の効果であると結論することができる⁵。評価研究の「黄金標準」(Gold Standard)と呼ばれる方法である。ただし、実験デザインによる評価事例が決して多くない社会科学分野では、それより低い質のエビデンスであっても利用せざるを得ない場合があるのが現状である(例えば回帰非連続デザインの適用など)。次に「2. そうした介入の事例の収集と体系化の作業」であるが、この作業は「システマティック・レビュー」(Systematic Review, SR)とも呼ばれる。複数の同種の評価結果を統合して、対象とした介入の効果があるのかわからないのかに関して一般的な結論を出す作業である。最後の「3. 広く利用されるような仕組の整備」であるが、これは、収集された事例や統合化して出された結論をデータベース化して、誰でもアクセスして参照できるようにしておくという作業である。

本特集では、医療、刑事司法、教育、ODAの分野における、上記の3つの作業の進捗状況について検証し、その検証結果を踏まえて日本での普及の可能性について議論した。

最後に、一つ注意を要することがある。それは、エビデンス（証拠）に基づく評価が各分野で注目を浴びていると言っても、それは世界的な動向であり、日本ではまだまだこれからの話だという点である。著者の知る限り、日本の社会科学の諸分野において、実験デザインを用いた評価はほとんど実施されていない。日本では、医療分野、農業分野、教育心理学、および工業生産管理の分野等の限られた分野で事例が報告されているだけである。したがって、レビューのプロトコルを定め、データベース化を進めると言っても、まずは大量の評価結果（一次研究）が産出される必要がある。各論文では、日本で大量の一次研究が産出される可能性とその道筋についても議論している。そして、この特集号自体が、エビデンス（証拠）に基づく評価に関する適切な認識の定着と、大量の一次研究の産出のきっかけとなることを期待している。

注記

- 1 本来、evidence-based という形容辞は、EBM (evidence-based medicine) に見られるように、エビデンスによって裏付けられる対象、つまり、その有効性が評価される対象に付加されるものである。しかし、evidence-based evaluationと言う場合のevaluationは、評価対象ではなく評価行為それ自体を指している。つまり、evidence-based evaluation とは、evaluation自体がエビデンスによって裏付けられていると言うことを意味するのではなく、さまざまな分野で進展している、エビデンス重視の流れに沿った評価活動の総体を意味している。本特集では、研究者の間で使用されているevidence-based evaluationという用語を使用してはいるが、本来は、evidence-based practices and policies (エビデンスに基づいた実践と政策)あるいは、evidence-based policy-making (エビデンスに基づいた政策決定)という用語の方が、実際のエビデンスの位置づけと利用目的を適切に表現していると言える。さらに正確を期せば、evidence-based よりも、evidence-centred、evidence-focused、あるいはevidence-informedと表現すべきだと言わざるを得ない。このように用語に関して多様な意見があるのが世界の評価研究の現状であるが、本冒頭文では、筆者（佐々木）の責任において、標記のタイトルを用いることにした。
- 2 あるいは逆に、確かに効果がないというエビデンスが確認された政府の介入行為を廃止するということも当然に含む。
- 3 連邦政府の改善を求める民間団体であるThe Council for Excellence in Governmentから資金援助を得た「エビデンスに基づいた政策を推進するための連合」(The Coalition for Evidence-based Policy)のウェブページから抜粋。
<http://coex.gov.securesites.net/index.php?keyword=a432fbc34d71c7>
- 4 社会科学では、無作為割当を用いたデザインを実験デザインというが、医療分野では当該デザインを「RCT: Randomized Controlled Trials」と呼んでいる。
- 5 実験デザインには、ここで紹介した最も単純なデザイン以外に、より高度なデザインも存在するが、無作為割当を適用するという基本的な考え方は同じである。

【研究論文：依頼原稿】

エビデンスに基づく医療（EBM）の系譜と方向性： 保健医療評価に果たすコクラン共同計画の役割と未来

正木 朋也

国際医療福祉大学大学院保健医療学専攻
masakit-ky@umin.ac.jp

津谷 喜一郎

東京大学大学院薬学系研究科
tsutani-ky@umin.ac.jp

要 約

国内外の公共投資について科学的なエビデンス¹を明らかにすべきとの社会要請に対し、日本における経験は保健医療分野を除いて皆無と言わざるを得ない²。しかし、イギリス、アメリカにおいては、公共政策判断の場においてもエビデンスに基づく評価が行われ、着実にエビデンスが蓄積されつつある。キャンベル共同計画、アメリカ教育省のWhat Works Clearinghouse、ODA分野の貧困アクションラボの事例などがそれにあたる。このような背景を踏まえ、本稿では、まず先行する「エビデンスに基づく医療」(Evidence-based Medicine: EBM)の系譜とコクラン共同計画について概観する。ついで、社会基盤としての質の高い情報の蓄積の重要性と社会認識の問題を踏まえ、EBMの今後の方向性について考察した。

キーワード

EBM、エビデンス、ランダム化比較試験、システマティック・レビュー、社会認識

はじめに

国内外の公共投資が効果的に実施されているかを考えるとき、プロジェクトやプログラムレベルにおける効果測定や判定も重要であるが、社会的コンテキストを踏まえたよりマクロな政策の評価もまた重要である。本稿では、このような視点から「エビデンスに基づく医療」(Evidence-based Medicine: EBM)と政策評価の系譜についての考察を試みる。まず、EBMという概念が誕生した歴史背景を概観し、その社会的要請がいかに達成されたか(されつつあるか)をみる。次いで、情報インフラの役割を担っているコクラン共同計画(The Cochrane Collaboration: CC)³について、シス

テマティック・レビュー(Systematic Review: SR)とその結果をまとめてデータベースとしたコクラン・ライブラリについて概説する。さらに、情報の質の確保について要点をレビューし、研究デザインと批判的吟味、研究結果の要約を効率よく伝えるための構造化抄録の重要性などについて述べる。

因みに、EBMは広義には、Evidence-based Healthcare、Evidence-based Public Health、Evidence-based Policy-making他、保健医療のみならず健康に関わるあらゆる領域を含む。一方、狭義には、文字通り、まさに一人の患者を対象とした医療現場での治療に関わる狭義のプラクティスを示す本来の意味としてのEBMがあり、その関

連書籍も多い。本稿では上述の視点から、後者の詳細については殆ど触れず、公共投資の評価とマクロレベルの政策に関わる視点からEBMを紹介させて頂くことをご承知頂きたい。本来、前者は広義のEBMとしてEvidence-based Healthcare (EBH) とでも表記し区別すべきものであるが、現実には、後者(狭義のEBM)をも含めた(合わさった)概念としてEBMという用語が普及・定着しつつある。EBMという用語を初めて目にされた方は、この状況に留意のうえで本稿をお読み頂ければ幸いです⁴。

このようにEBMには、広義と狭義の意味があるが、その根底には対象となる患者の好み(patient preference)、価値システム(value systems)、および現場の状況などをも踏まえた共通性がある。時に「ランダム化比較試験(randomized controlled trials: RCT)に基づく医療」(RCT-based Medicine)との批判を受けることもあるが、その認識は誤りである。EBMの基本は、研究によるエビデンスの産出にあり、そのエビデンスにはグレードがあることを明白にしたうえで、それを場の状況に応じて適切に利用することである。

実績と経験にも支えられたEBMの考え方は極めて有益であるので、他の分野・領域においても、その現状に則した評価活動に活かすことが可能である。

本稿ではさらに、この基本的な考え方を支える科学の公衆的理解(public understanding of science: PUS)の重要性にも触れ、適切なナレッジ・マネジメントの仕組みが、基本的な社会要請(事実を正確に公表した上で公明正大な判断を行う)に対するひとつの有力な回答となり得ていることを述べ、社会認識(social epistemology)に及ぼす影響と未来展望についても言及した。

1. 政策評価とEBM

Evidence-based Medicine (EBM) という用語は、カナダのマクマスター大学のGuyattの論文(Guyatt 1991)に用いられたのが始まりとされる。EBMという用語はその後、特にSackettらの“Evidence-based Medicine: how to practice and teach EBM”

(Sackett 1996)発行後、さらに世界的に用いられるようになった。

日本では、「科学的根拠に基づいた医療」などと訳されることもあった。上述の広義・狭義の違いに加え、「科学的根拠」という一般的用語として理解されることによる議論が生じ易いため、昨今著者らは「エビデンス」というカナ標記を用いることを推奨している。

ここでは、政策評価の変遷を振り返りながら、EBMがどのような時代背景に生まれ育ったかについて歴史背景をもとに考察する。

(1) イギリスの研究開発戦略とエビデンスに基づく保健医療(EBH)

イギリスでは1980年代の行政改革の動向を背景に、医学研究に対する深刻な論議が起こった(Smith 1988)。当時の論点は、研究活動の低下、基礎研究と応用研究のバランスの崩れ、研究結果の未利用、であった。上院が委員会を設置し問題を検討した結果、健康サービス研究(health service research)と公衆衛生研究(public health research)を促進すべきという結論に至った。当時としては、大変驚くべき結論であったが、政府はこれを歓迎した。

こうした研究の方向性の転換は、英国保健省(National Health Service: NHS)のサービス提供と国家的な研究投資の費用-効率に大きな影響をもたらした。1990年に政府はNHSの中核に研究開発部門を設置して1991年には保健医療の研究開発戦略を打ち出し、国民の健康・福祉の改善、科学的な情報に基づく医療政策と医療行為の促進、効果的で効率的な資源の利用、のための行政改革を実施した。これらの改革は、それまでの意思決定者の意見に基づく政策決定から、エビデンスに基づいた政策決定へとシフトするきっかけ作りになったとともに、政策実施により成果を挙げたかどうか、納税者である国民に対しその説明責任を果たす試みでもあった(ミュア・グレイ1999⁵)。この改革の目的は、エビデンスに基づく保健医療(Evidence-based Healthcare: EBH⁶)そのものといえる。

このような背景の中、1992年10月に研究開発部

門の情報ネットワークと協力する形でThe Cochrane Centreがオックスフォードで活動を開始し、翌1993年3月にThe UK Cochrane Centreと名称変更の後、10月の第1回コクラン・コロキウム（Oxford）により正式にCCとしての活動を開始した。また同年にはレビュー作成と共にイギリスの医療関係者をはじめとして、情報を必要とする人々にレビューとその結果の意味を普及させる目的で、政府直轄のNHSレビュー普及センター（NHS, the Centre for Reviews and Dissemination: NHS CRD）がヨーク大学内に設置された。

イギリス政府の行政改革や世論としての研究開発思考の追い風もあり、計画的かつ継続的に質の高いエビデンスを維持・更新する基盤が確立した。アメリカの研究者や行政評価の立役者らとも相互に呼応して、その後、十数年にわたる情報と経験の蓄積を経て、ようやく今日のゆるぎない保健医療技術評価基盤を形成し得たものと理解できる。広義の概念を擁するEBHの発展は、それぞれ保健制度の異なるイギリスとアメリカにおいてほぼ同時に進展し、情報技術のめざましい発展の恩恵も受けながら、次第に世界中で受け入れられ、多くのボランティアらの協力を得て現在もなお拡大展開しつつ今日に至っている。

（2）アメリカを中心とした政策評価の変遷とEBM

アメリカの政策評価の歴史を振り返ると1930代の教育と公衆衛生の分野にその原型をみることができる（Rossi et al 1999）。1965年合衆国連邦政府は、ジョンソン大統領が打ち出した2つの政策的スローガン（貧困との戦い、偉大なる社会の建設）達成のため全ての省庁に予算配分システムPPBS（Planning Programming Budgeting System）を導入し、事業目的、成果評価、代替案の作成、効率性評価という一連の手続きの下に予算配分を行った。このPPBSの影響で、これまであまり脚光を浴びることの無かった政策評価についても、その評価手法や方法論に関する研究と実績が飛躍的に増えることとなった。なお、PPBSは1960年代後半に合衆国連邦政府で導入が開始され、地方政府も含めて一応の定着を見たプログラム評価

（今日的な意味での政策評価）の起源ともいえるべきものであったが、その定着は失敗したとの見解がある（窪田1998）。

その後、1980年代に入り、財政赤字の抑制を目的とした公共プログラムの絞込みと効率化のための道具として政策評価が採用されるようになり、その結果この時期には、評価手法・評価研究それ自体の質的な向上と発展がみられた。

1990年代は市場経済の波に押され、政策評価についても財務的・経済的な効率性の比重が高くなり、科学的な評価からより実用的な評価へと評価対象が広がっていった。連邦政府は1992年に行政改革の綱領ともいえるNational Performance Review（NPR）を発表し、より少ないコストで、よりよい成果を上げる、消費者を最優先に、手続き重視から結果重視へ、を掲げ一般に、より結果が理解され易い評価を重視する機運が醸成された。1993年には「政府業績成果法」（Government Performance and Results Act: GPRA）も成立し、事前・事後の比較等を応用し、計量可能な指標を用いたパフォーマンス・メジャーメントによる評価が広範に利用され始めるに至った（龍、佐々木2000、田辺2003）。

このような背景の中、連邦政府は医療分野に関してもGPRAの下位に保健医療法（Health Care Act）を制定し、NPRの具体的項目を達成するために必要な法的整備を行った。NPRでは行政施策が何を目指し、どのような手続きを経て、どういう目的を達成したかをみることにしていた。その手法は、Evidence-based Management、Result Oriented Management、Performance-based Management、Managing for Resultsなどさまざまな用語で呼ばれているが、何れも経営管理論の始祖として知られるアンリ・ファヨール（Henri Fayol, 1841-1925）の示した“Plan, Do, See”サイクルの循環の上になり立つ。即ち、事前に定めた計画に従って、目的に合致した活動を実行し、その効果を測定し評価するという、改善に向けての基本的かつ連続的なアプローチであった。

John Rodman Paul（1893-1971）が1938年に米国臨床研究学会で提唱し（Paul 1938）、Alvan R. Feinstein（1925-2001）により始められた臨床疫学（Feinstein 1968）におけるOutcome Researchや

Decision Analysis、Cost-Effectiveness Analysisなどは、上述のマクロレベルの社会改革の進展に呼応した医療評価におけるアプローチといえる (Sackett 1991)。その背景には、経営管理医療 (マネジド・ケア) や健康維持組織 (health maintenance organization: HMO) による医療の効率化とNPRの医療適用について、当時北米を席卷していた社会改革の手法が少なからず影響したであろうことは容易に想像できる。

重要な点はEBMそれ自体が独立先行し現在に至ったわけではなく、むしろ社会改革の一翼を担い、また、社会の要請に応える形で、他の社会改革評価手法と共鳴しつつ発生・発達し得たことである (林2000)。

2. コクラン共同計画とシステマティック・レビュー

(1) エビデンスを「つくる」「つたえる」「つかう」

コクラン共同計画 (The Cochrane Collaboration: CC) は1992年にイギリスの国民保健サービスの一環として始まり、現在世界的に展開されている医療技術評価のプロジェクトである。ランダム化比較試験 (randomized controlled trial: RCT) を中心に、世界中の臨床試験のシステマティック・レビュー (SR, 収集し、質評価を行い、統計学的に統合する) を行い、その結果を、医療関係者や医療政策決定者、さらには消費者に届け、合理的な意思決定に供することを目的とし、EBMの情報インフラストラクチャーとしての役割を担っている。

コクランというのはアーチボルト・コクラン (Archiebald Cochrane, 1909-1988) という医師で疫学者の名であり、彼は1979年に「各トピック毎に、それぞれのランダム化比較試験を、定期的にクリティカルにまとめていないことに関して、われわれ専門家は批判されるべきである。」と述べた。

この意味するところは、すでにあるRCTから、よいものを選びすぐり質の悪いものは省いて、それらをまとめて、遅滞なく、必要な人に届ける、ことが大切だということである。

エビデンスは「つくる」「つたえる」「つかう」ことに意義があり、それに関わる立場も3つある (表1)。コクランの発言で最も重要な点は、エビデンスを「つくる」場と「つかう」場の間に入って「つたえる」ことが大事だと力説したことである。なお、エビデンスは「つかって」こそ意味があるのだが、実際にはこの「つたえる」ことは意外に容易でないことについては後で述べる。

表1 エビデンスに関わる3つの立場

つくる	臨床試験
つたえる	コクラン共同計画
つかう	臨床家、行政官、消費者

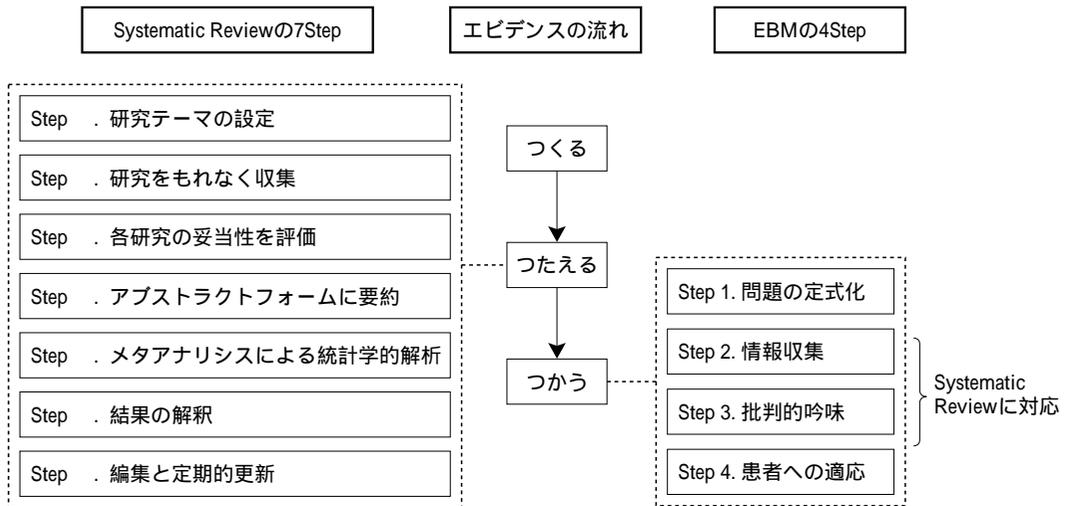
(出所) 津谷2000a

(2) システマティック・レビュー (収集し、質評価を行い、統合する)

CCはコクランの弟子のIan Chalmers というイギリスの産婦人科医が始めたものである。彼はコクラン共同計画を始める前から、10年以上にわたり周産期領域のシステマティック・レビュー (SR) を行っていた。CCの「つたえる」ことの前半部分 (情報の生産) の中心的作業がSRである。後半部分の作業は利用者へのサービス提供であり、そのプロダクトが後に述べるコクラン・ライブラリである。

SRの具体的手順は図1のステップによるが、手順の詳細とCCが提供するシステマティック・レビュー・ハンドブックなどの概要については既報 (津谷1997, 2000a, 2000b) を参照されたい。ここで強調しておきたいことは、部品としてのRCTを実施すること (あるいはエビデンスをつくること) 自体はもちろん重要であるが、既存の蓄積された研究成果 (領域によっては必ずしもRCTがなされているとは限らない) を、批判的・系統的に吟味し、重要な結果のみを抽出・整理・上清してゆくSRのプロセスこそが最も重要であるという点である。即ち、SRはCCをはじめ、キャンベル共同計画、アメリカ教育省のWhat Works Clearinghouse および貧困アクションラボなどに共通する活動の本質ともいえる。

図1 システマティック・レビュー



(出所) <http://www.lifescience.jp/ebm/opinion/200308/#fig01> (accessed Jan 3, 2006)

表2 コクラン・ライブラリ

Database	Total Records
1) The Cochrane Database of Systematic Reviews (Cochrane Reviews)	4,113
2) Database of Abstracts of Reviews of Effects (DARE)	5,585
3) The Cochrane Central Register of Controlled Trials (CENTRAL)	463,763
4) The Cochrane Database of Methodology Reviews (Methodology Reviews)	20
5) The Cochrane Methodology Register (Methodology Register)	7,484
6) Health Technology Assessment Database (HTA)	4,733
7) NHS Economic Evaluation Database (NHS EED)	16,418
8) About The Cochrane Collaboration and the Cochrane Collaborative Review Groups (About)	90

(出所) The Cochrane Library, <http://www.thecochranelibrary.com/> (accessed Jan 3, 2006)

なお、コクラン共同計画では上記全体のプロセスそのものをシステマティック・レビューと呼び、統計学的・数学的なプロセスの部分をメタ・アナリシスと呼んでいるが、一般には全体をメタ・アナリシスと呼んでも差し支えないと思われる。昨今では、時に、質的研究のメタ・アナリシスという言葉も耳にするが、この場合は既存研究を統合する研究の総称であり、さらに広い意味でメタ・アナリシスという用語が用いられている。EBM同様、狭義と広義の意味の違いには注意が必要である。

(3) コクラン・ライブラリ

先にCCの主要な活動の「つたえる」ことには、前半部分（情報の生産）と後半部分（利用者へのサービス提供）があることを述べた。その後半部分のプロダクトがコクラン・ライブラリである。CCでは、上述のSRを経て四半期に一度コクラン・ライブラリを更新し有料でその成果を一般に還元している。CD-ROMに専用の検索ソフトをバンドルし販売するとともに、ホームページからWebによる利用も可能である。日本の医学系学部をもつ多くの大学で利用可能であり、個人などで

の購入も可能である⁹。

コクラン・ライブラリは8種類のデータベースからなり、2006年1月時点でそれぞれ表2に示す件数の登録がある（The Cochrane Library 2005, Issue 4）。

それぞれのデータベースの内容は以下のとおり。

1) The Cochrane Database of Systematic Reviews (Cochrane Reviews)

コクラン・ライブラリの中心的データベース。各種疾患を対象とした治療や診断についてのSRを収録し、Complete reviewsとProtocols（レビュー中のもの）からなる。CCのCollaborative Review Group（CRG）が“Cochrane Reviewers' Handbook”に基づき作成。

2) Database of Abstracts of Reviews of Effectiveness (DARE)

学術誌等で公表されたSRを収録。イギリスのヨーク大学NHS Center for Reviews and Dissemination（CRD）において、基準を満たすレビューを選択し、構造化抄録を作成。

3) The Cochrane Central Register of Controlled Trials (CENTRAL)

RCT検索の中心的データベース。ランダム化比較試験や比較臨床試験（準ランダム化）の論文、学会抄録等の書誌情報を収録。非公開の臨床試験も収録している。

4) The Cochrane Database of Methodology Reviews (Methodology Reviews)

SRの方法論を検討しているMethodology Review Group作成のデータベース。

5) The Cochrane Methodology Register (Methodology Register)

SRに関する方法論（作成手法・ランダム化比較試験の研究手法）論文の書誌情報のデータベース（EBM実践の為に資料集）。

6) Health Technology Assessment Database (HTA)

医療技術評価の文献を収録。治療、予防、リハビリテーション、医療機器、手術手法などを対象。NHS CRDにより維持管理。

7) NHS Economic Evaluation Database (NHS EED)

医療の経済効果、費用分析論文の構造化抄録・書誌情報データベース。NHS CRDにより維持管理。

8) About the Cochrane Collaboration and the Cochrane Collaborative Review Groups (About) コクラン共同計画そのものの情報。

表2上の8) About The Cochrane Collaboration and the Cochrane Collaborative Review Groups (About) をみると、ミッション・ステートメントとともに、各運営委員会メンバーが過去5年以内にどこから資金援助を受けたか、コンサルテーションに関わる報酬を得たか、などの情報が公開されており、独立性と中立性を自ら示そうとする姿勢も興味深い。

また、Collaborative Review Groups (CRGs) のリストもあり、終了したReviewや現在進行中のProtocolの全てを一覧することができる。

なお、DARE、HTA、NHS EED は、NHS CRD のDAREサイト⁹にて無料で検索可能である。

Cochrane Reviewのabstractについても無料でweb上で公開されている。その日本語訳は日本のネットワークであるJANCOGでボランティアになされていたが、作業量の多さなどから中断した。その後、(財)日本医療機能評価機構のMINDSで診療ガイドラインに関連したCochrane Reviewを中心に訳がなされ、2006年3月から無料公開がスタートした。

(4) 組織体制、レビュー・グループ、関連ネットワーク

CCは独立した中立機関として活動を展開し、現在12のセンター（各支部を加えると計25）が地域毎の教育や活動支援を行うとともに、レビューの重複が生じないように調整と維持管理にあたっている。

これらのセンターを拠点に、現在50のレビュー・グループが活動しておりその殆どが独自のホームページを持ち、進行中のレビュー内容について相互に参照できるようにしている。レビュー・

グループの殆どは疾病・治療領域毎に形成されているが、中にはMethodology Review Groupのように、レビュー方法そのものについての検討グループもある。

CCではさらに、異なる研究グループとの協力体制をとるため13の分野（カウンターパート/専門グループを含む）について担当者を定め公表している。キャンベル共同計画や各領域の専門家グループらと相互に連携することにより、常に最新の動向を共有するとともに、重複したレビューの調整などを行うことが目的である。同様の相互連携や協力体制はキャンベル共同計画の教育グループとアメリカ教育省のWhat Works Clearinghouseとの間にもあり、目的と方法論を共有できるグループ同士が相互の活動を尊重しつつ連携をとることは、SRを推進する立場や過去の歴史的な背景から考えても極めて自然な流れといえる。

3. 情報の質の確保

インターネットの世界的な普及により、情報は指数関数的に増加している。このように情報が氾濫する中で、我々はどうのように情報を取捨選択し、合理的な判断を行うことができるだろうか？このような、情報の質の確保とその評価は、行政判断あるいは個人の生活に直接影響する問題であるにも拘わらず、多くの場合なおざりにされている。情報が日々刻々と変容することを考えれば、この問題は受身の姿勢では解決できないことは明らかである。ここでは、情報の質をいかに積極的に高く維持するかという観点から、エンドポイント、バイアス、試験デザイン、エビデンスのレベルについて留意事項を述べる。

(1) エンドポイント (endpoint)

臨床研究において用いるエンドポイントという用語は、評価目標、評価項目、評価基準、評価尺度、評価のものさし、といった意味であり（津谷1995）ICH統計ガイドライン（E9）¹⁰の定訳として「主要変数（「目標」変数、主要評価項目ともいう）」がある。英語ではindicator、marker、

parameterなどと呼ばれるもので、臨床評価や薬効評価においては、時に、主要評価項目（primary endpoint）副次的評価項目（secondary endpoint (s)）などにより介入効果を測定し、効果判定のよりどころとする。要は当該研究で「何を見たいか」をエンドポイントとし、一番重要な項目を主要評価項目と定義して重要視し、それ以外のものを副次的評価項目として参考とする。

主要評価項目が疾病による生死といったように観察しづらい場合には、その代替エンドポイント（surrogate endpoint）を用いる場合もある。これは真のエンドポイント（true endpoint）に対する用語である。例えば糖尿病や高脂血症の患者を治療対象として考えるとき、本当は薬剤投与の効果による延命（生死が真のエンドポイント）を臨床研究の結果としてみたいのであるが、倫理的、経済的、および時間的制約などの実施可能性から生死を主要評価項目とすることができないため、過去の研究などから疾患の病態と関連することが明らかになっている血糖値やコレステロール値を代替エンドポイントとして臨床試験を実施し、本来知りたいエビデンスを間接的に産出する。なお、代替エンドポイントが正しく選択されていたかについては、通常、平行あるいはその後実施される疫学調査などにより、後になって明らかにされることになる。

エンドポイントは「客観性objectivity」の高いものとすべきとの考え方が強く、数値や画像で表されるものや、死亡のように「あり」「なし」で表されるものは明確にみえるので、ハード・データと呼ぶ。一方、生活の質（Quality of Life: QOL）のように本人の「主観性subjectivity」や時に「第三者の主観性」を重視したものをソフト・データと呼び区別する場合もある。

何れの場合でも、エンドポイントは客観的であることが望まれる。しかし、「主観性」はもちろんであるが「客観性」にも社会文化的な価値観が反映されるため「客観性」のあるエンドポイントを設定することは必ずしも容易ではない。特に文化的価値観に相違のある歴史的あるいは国際的な研究を行う場合には「客観性」という言葉が常に曖昧さを含んでいることには十分留意する必要がある。

なお、「信頼性reliability」と「再現性reproducibility」は結果的にほぼ同義と考えることができる。観察者や測定日時が変わった場合でも同様の結果を得るために「信頼性（再現性）」の確保が重要である。「信頼性（再現性）」が確保されていればその結果の評価は容易になる。

しかし、次に述べるような、様々な原因により発生するバイアス（bias）に妨げられて再現性が確保できないことは評価の現場ではよく経験されることである。以上を踏まえれば、科学的なエビデンスを産出するためには上述した問題を十分意識したうえで、積極的にこれらバイアスを排除しなければならないことが改めて理解できる。

（2）バイアスと研究の事前登録

あらゆる研究や評価には必ず何らかのバイアスが入り込むと考えるべきである。統計学では傾向のないランダム誤差（random error）であるバラツキ（variability）に対して、バイアスは系統誤差（systematic error）として捉えられる。

マクマスター大学のハダットの著書（ハダット2004¹¹）には36種類ものバイアスが紹介されている。1982年に報告されたもの（Owen 1982）を補足し「利用者が情報を理解する過程で生じるバイアス」として紹介している。ここで各バイアスの詳細説明は省略するが以下の括弧内の語の後にバイアスという語をつけ、語感から想定されるバイアスの内容を推察すると面白い。

「競争」「彼には借りがある」「モラル」「専門領域」「何かをしなければならない」「好みのデザイン」「有名雑誌」「有名な著者」「大規模試験」「はでなタイトル」「資格あるいは専門的なバックグラウンド」「地理的」「公表物の言語による」「怠慢」「伝統」「銀行通帳」「反抗」「テクノロジー」「経験主義」「私は疫学者である」

有名の反対には無名、大規模の逆に小規模といった表裏をなすバイアスがあることを考えれば、さらに多くのバイアスが存在することが理解できる。バイアスというと、研究デザインを立案するときに可能な限り取り除くべきものであるといっ

たことを考えがちであるが、得られた結果を解釈する際にも気づかぬうちに多くのバイアスを自ら抱え込んでいることは認識しておくべきであろう。

また、研究者やスポンサー側からみて良い（あるいは悪い）結果だけを報告するといった出版バイアスにも注意が必要である。こうした問題に敢えて立ち向かうことを表明し2002年に発行を開始した基礎医学の電子雑誌Journal of Negative Results in Biomedicine¹²は、出版バイアスが研究全体の方向性に誤った示唆を与え、ひいては社会認識に問題を生じさせることに対する警鐘を与えている。

2004年9月、国際医学雑誌編集者委員会（International Committee of Medical Journal Editors: ICMJE）は、臨床試験の事前登録がなされていることを雑誌投稿の条件とした。翌月カナダ・オタワにて、国際的な臨床試験の登録をサポートする原則を構築し、さらにその実施を推進することを目的とした会議が催された。これを受け、世界保健機関（WHO）は、2005年にInternational Clinical Trials Registry Platform（ICTRP）¹³プロジェクトを立ち上げた。これは世界に多数存在する臨床試験登録システムを、一定基準の下に束ね、これらの臨床試験登録システム内で一意となる試験識別番号を発行することにより、世界中に散在する臨床試験情報を統一的に検索できるようにすることを目指した活動である。

日本においては、2005年6月に「UMIN臨床試験登録システム」（UMIN Clinical Trial Registry: UMIN-CTR）¹⁴が立ち上がり、2006年1月時点でICMJEの了承を得るに至った。また時を同じくして、財団法人日本医療情報センター（Japan Pharmaceutical Information Center: JAPIC）¹⁵も臨床試験情報の登録と開示に関する情報提供環境を整備し、広く一般にその情報を公開し2005年7月より運用を開始している。さらに、日本医師会活験促進センター¹⁶において、医師主導活験にかかわる情報が提供されている。

いずれも登録は任意であり、開示する内容についても、登録者側に委ねられている。なおその際、治験¹⁷の登録において、治験薬コード等を含んだ情報を登録・公開した場合、薬事法上の未承認薬

の広告規制に抵触するのではないかと議論も生じたが、厚生労働省は2005年6月、「直ちに広告には該当しない」旨の見解を示した。これら一連の背景を踏まえ、日本製薬工業協会医薬品評価委員会では「臨床試験の登録・結果公開に関する実施要領」を定めるなど、国内製薬企業においても具体的な運用が始まっている¹⁸。因みに、このような臨床試験登録については、CCとの協力のもとにGlaxoWellcome社が1998年に先陣を切って始めていたことは特筆に価する (Sykes 1998)¹⁹。

エビデンスを正しく伝えるためには、研究の独立性や中立性を積極的に示すことも必要である。例えば、喫煙に関する研究に際し、たばこ産業から支援を受けたか否かを明らかにすることは必要であろう。また特定の団体や企業から支援を受けている研究者に対しても同様で、その団体や企業から支援を受けていたか否かについて公表することにより、他の報告書も含め、研究結果に潜在するバイアスの可能性の有無についての情報を補足することになる。国内では2005年に施行された個人情報保護法との兼ね合いで容易に実現できない可能性もあるが、既にこうした試みは、米食品医薬品局 (Food and Drug Administration: FDA) においては実施されている。

(3) 研究デザインの基礎、ランダム化比較試験 (RCT)

紙面の関係で、各種研究デザインの詳細については他の書籍に譲るが、ここではRCTの要点をレビューするとともに、RCTが保健医療領域でここまで普及した背景について、またその他の領域であまり普及しなかった点について考察する。

上述のバイアスを避ける唯一の方法がRCTのデザインであり、介入 (intervention) の真の効果や因果の推論を行ううえで極めて重要である。仮に未知のバイアスが入り込んだとしても、ランダム化により、そのバイアスを均等に振り分けることができ、集団として同質の背景 (性質) をもつグループ分けが可能となる。この同質の集団に対して、特定の介入を施しその結果を比較することにより、介入の効果のみを純粋に抽出できるところがRCTの研究デザインの本質である。自然変動や

想定しない内外の影響があるため、通常、単純な前後比較から、介入効果のみを抽出することは困難である。

ときに、RCTを「無作為抽出試験」と誤訳されることがある。これは想定する母集団から無作為 (ランダム) に被験者を抽出するプロセス「無作為抽出」(random sampling) と、抽出された被験者を二群にランダムに振り分けるプロセス「無作為割付」(random allocation) が混同されたことによる。臨床試験は、抽出したサンプルから得た結果を母集団に対して一般化 (generalize) しようという試みであり、その際、前者の無作為抽出は試験結果の「外的妥当性」を、後者の「無作為割付」は「内的妥当性」を保障するものである。

臨床試験の場でRCTというと通常ブラインド化 (blinding²⁰またはマスキングmasking) が前提となる。Concealment (隠蔽)²¹とマスキングとはよく混同されるが、前者は割付方法や割付順を、介入を行う前に介入を与える人に知られないようにすることである。一方、マスキングは介入が始まった後で患者や医師、結果の評価・解析者などに割付内容を知らせないことである。隠す時点と隠されている対象者が異なる。二重盲検のデザインや中央割付方式によるランダム割付が実施されている場合には、通常concealmentも確保されていると理解できる。

マスキングしない状態で数種類の薬剤を割り付けることを考えると、例えば、現場の医師であれば、割付対象となった重症患者を前に低用量群やプラセボ (偽薬) を割付することに倫理的な問題を感じ、割付指示どおりの投薬をしないという行動 (バイアス) が生じる。このような非ブラインド、あるいは単ブラインドの研究デザインの問題点は臨床研究を重ねるうちに経験的に把握されたもので、現在保健医療領域においては、それらの方法から質の高いエビデンスを得にくいとの認識が一般的である。この領域でRCTといえば暗黙の了解の下、二重マスキング (かつランダム化比較) 試験を示すと理解される²²。二重とは治療を施される患者と治療を施す医師の両者がともに、割り付けられた治療方法の中身を知らないという意味である。さらに、必ずしも一般的ではないが、評価者・統計解析担当者にも治療法をマスクする

ことにより、評価・解析時のバイアスを取り除く三重・四重ブラインドのデザインもある。

臨床試験には、評価を目的とした、ヒトを用いて、意図的に開始される、科学実験、といった四つの視点がある。ヒトを対象として意図的に行うため、当然倫理的な問題が生じる。臨床研究においてはその必要性から、こうした問題に正面から取り組み、実施にあたっての様々な手続きや法制化²³によりこれらを解決してきた。例えば、インフォームド・コンセント (informed consent) の下に自由意志で試験に参加することや、事後の同意撤回は今や常識であり、後で同意を撤回してもその後の治療に不利益を受けないことなどについては、臨床試験にかかわる全ての人で合意形成がなされている。

未知の副作用などの負の作用の可能性もあり、必ずしも新しい治療法が優れているとは限らないのであるが、新薬の方が既存薬よりも優れているという一般認識がある。割付の結果、既存薬や比較対照として用いるプラセボ投与を受けることになる可能性についても、事前に説明し合意を得るが、ときに被験者側からどうしても新薬の治療を受けたいといった希望もありうる。このような場合には、クロスオーバー法²⁴により、二種類の治療を時間的に順番を変えて交互に全員に施すことにより参加者への不公平感を是正することも可能である。通常この方法は臨床試験の初期段階において少数の被験者から多くの情報を引き出すために利用されているが、不公平感の是正という観点から教育分野や国際援助の現場でも応用可能である。

臨床研究で経験した様々な問題は、同じく「人」を対象とする人文・社会科学の研究においても同様に生じ得る。保健医療とそれ以外の領域におい

てRCTとSRが普及しない理由のひとつに、これらの手間のかかる問題を克服してまでRCTによるエビデンスを産出する必要性と合理性が認識されない (他の手法で十分であると考えている) ことが推察できる。また一方で、保健医療でエビデンスの集積が可能となった背景には、健康問題は行政も個人も共に直面する強い興味対象であったことがあげられる。さらに、イギリスNHSの研究開発戦略の追い風も受け、適時に予算確保が可能であったことは保健医療領域でのエビデンスの蓄積に大きく貢献したものとと思われる。保健医療以外の (場合によってはむしろ予算削減された) 領域は、これらの点で結果的に大きく水を開けられることになったものと考えられる。

なお、保健医療領域外のRCTについては、実施上の制約などからマスキングしない (できない) 状態で試験を実施することの方が多くようである。RCTの品質を読者が評価する場合にも、評価がマスキングされた状態でなされたか、あるいはオープンな状態でなされたか、に依存する系統的な相違が存在することが報告されており (McNutt 1990) 今後RCTの経験が増えるにつれ、保健医療領域で経験した同様の問題が浮上する可能性はある。なお、遺伝情報など個人の利害に関わる情報を扱う場合には、倫理的な観点からその個人が特定できない配慮を施すことは当然であるが、マスキングと患者情報の秘匿とは独立した異なる問題である。

(4) エビデンスの質とお勧め度のグレーディング

エビデンスとは、「バイアスのない方法により得られたデータを、バイアスのない方法で分析して得られた結果」の総称であり、介入結果のほか、

表3 エビデンスのレベルと内容

a	ランダム化比較試験のメタ・アナリシスによる
b	少なくとも一つのランダム化比較試験による
a	少なくとも一つのよくデザインされた非ランダム化比較試験による
b	少なくとも一つのほかのタイプのよくデザインされた準実験的研究による 比較試験や相関研究、ケース・コントロール研究など、よくデザインされた非実験的、記述的研究による 専門家委員会の報告や意見、あるいは権威者の経験

診断や予後に関わる結果についても産出され得る。そのレベルは科学的研究結果の質や水準に依存し、その研究デザインによって、通常、メタ・アナリシス>ランダム化比較試験（RCT）>比較臨床試験（CCT）>コホート研究>症例対照試験（case-control study）>症例集積研究（case series）>1例報告（case report）の順番にレベルが高いとされる。以下に示す、米国の医療政策研究局（Agency for Health Care Policy and Research: AHCPR, 1993）のグレーディング・スケールもこのような考えによっている（表3）。このスケールの基本的な考え方はその後多くのバリエーションを生んだ。その代表的なものがOxford Centre for Evidence-based Medicine²¹の分類（2001）として取り入れられている。

EBMにおいては通常、RCT以上のレベルの臨床情報をエビデンスが高いものとして扱うが、非ランダム化あるいはその他の研究スタイルであっても、よくデザインされた研究の結果は、専門家や権威者の報告や意見よりも高いレベルのエビデンスとしている点に注目すべきである。別の見方をすれば、領域により、RCTの実施が困難な場合であっても、よくデザインされた研究を実施すれば当該領域内においてエビデンスを「つくる」ことが可能であることを示している。重要なことは領域毎に異なるエビデンスのレベルを提供者と利用者間で正確に共有することであろう。このようなレベルの概念を含むエビデンスを精確に表現し伝える努力が必要であるが、最終的にはどうしても「つかう」側の認識の問題が介在するため、うまく「つたえる」ための方策や施策が重要となる。

なお、上述したように複数のグレーディング・スケールが並存する中でその運用についての欠点と問題点を改善する目的で有志が集まり、2000年よりGRADE Working Groupとして非公式な活動を開始している²⁶。彼らは、2004年にはエビデンス利用者の視点に立ち、「エビデンスの質」とそのお勧め度（strength of recommendation）を評価するシステムを開発した（GRADE Working Group 2004, 津谷他訳 2005²⁷）。その中で、個々の主要なアウトカムにおけるエビデンスの質は、(1) 研究デザイン、(2) 研究の質、(3) 一貫性、(4)

直接性、をそれぞれ考慮したうえで決定すべきとしている。そこでは、まず研究デザインをRCTと観察研究（コホート研究、ケース・コントロール研究、断続的な時系列研究、コントロールされた前後比較研究）に分類し、その介入による利益が害より多いか否かを利害得失の4段階（正味の利益あり、利害得失あり、不確かな利害得失あり、正味の利益なし）に分類することを提案した。さらに、これらをもとに「お勧め度」のグレード4段階、しなさい（Do it）、した方がよい（Probably do it）、しない方がよい（Probably don't do it）、するな（Don't do it）のカテゴリーを提案した。

彼らの活動の重要な点のひとつは、意思決定に必要な簡潔性（simplicity）と明瞭性（clarity）をうまくバランスするよう試みた点である。システムを単純化すれば判断が黙示的になりやすく、明瞭性を欠くことになり、逆に、明瞭性を高め意思決定のプロセスをわかりやすくしようとすれば、システムは複雑なものとなりやすい。もうひとつの重要な点は、彼らのアプローチは、構造化された熟考（structured reflection）のための枠組みを提供するものであり、適切な判断を確実に行う助けとなるが、ユーザー自身の判断の必要そのものをなくしてしまうものではないとする立場を明確にし評価システムを開発・提案している点である。

(5) 構造化抄録のすすめ

エビデンスを効率的に「つたえる」ための一つの方策として構造化抄録がある。これは学術雑誌編集者らにより提案されたもので、介入研究の場合に限られるがRCTとメタ・アナリシスの論文・報告書の書き方については1996年にCONSORT声明^{28,29}としてまとめられ、国内でも同様の構造を用いた記述が推奨されている（青木2000）。中山らは欧文医学雑誌の構造化抄録書式について検討し、従来から一般に科学論文で用いられてきた抄録記述のIMRAD構造の書式（Introduction, Method, Results And Discussionの頭文字）と実際に用いられていた8項目の対応を示した（Nakayama 2005）。(1) ObjectiveはIntroduction、(2) Design、

(3) Setting、(4) Patients or participants、(5) Interventions、(6) Main outcome measurementの5項目はMethod、(7) ResultsはResults、(8) ConclusionsはDiscussionに対応したと報告している。

ここで重要な点は、で研究の目的を明確に記述すること、追従する研究者らが追試を再現できるよう方法については～を漏れなく書くことである。また、の結果をもとにした単なる議論よりも、で、良かったか、悪かったか、どちらでもないか、といった結論をきちんと書くことが重要である。これらが押さえられていれば、読者は論文の本文を読む前に、研究結果とその結論を容易に知ることができ、その上で必要に応じて試験デザインや目的を知るといった、効率的な情報共有が可能となる。また研究報告者側としても、これらの抄録を書くために、研究開始前から試験目的やデザインを明確にしたうえで、どういう結論を述べるかを事前に十分検討する必要が生じることになり好ましい。さらに、論文化する時点においても標準的な基本構成に従って記述することから、重要な情報の記載漏れを防ぐことにも役立つ。

研究論文の報告形式として構造化抄録が世界的に広く認識されつつある一方で、国内学術雑誌の取り組みが遅れていることも明らかにされた^{30,31,32}。これは、EBMに基づくインフラが国内においてははまだ整備できておらず「つたえる」ことの重要性が十分認識されていないことを示唆する状況である。エビデンスを「つたえる」ところは、コクラン共同計画などの活動だけに委ねるのではなく、エビデンスを生産する研究者も、それを届ける出版社などとともに情報提供環境の整備に貢献できることを再認識すべきであろう。

4. 情報の共有化と社会認識

エビデンスに基づく健康施策や政策を考えると、SRの結果、整理・蓄積された情報を、どのように政策に反映して利用できるかは社会認識 (social epistemology) の視点を含めて考える必要がある。公衆の健康を考える場合蓄積されたエビ

デンスは貴重で有益な情報を与えるが、施策や政策が他の諸々の情報をもとに決定されることを踏まえれば、いかに明白なエビデンスであろうとも、ひとつの要素に過ぎない。エビデンスを政策決定へ統合させ、広く社会に還元させるために、どのような問題と解決策があるか考察した。

(1) 説明のジレンマ

マスメディアによる(エビデンスや保障のない)興味本位の報道の方が容易に社会に伝播・浸透する現実があり、ときに、重要なエビデンス情報を伝えることの弊害にもなり得る。これをうまく回避する方法論が必要になるが、一般に、複雑な情報を専門家でない人に説明する場合、細かく説明すればするほど(情報量が増せば増すほど)かえって本質の理解から遠ざかるということはしばしば経験される。いわば「説明のジレンマ」とも言うべき障壁である。個人の理解を促す最適な情報量というものは伝達する情報毎に必ずどこかに存在するはずであるが、それは個人の背景知識や認知能力にも依存し、また、質的にも全く異なるものと理解すべきである。従って、「説明のジレンマ」を前に、個人に適切かつ正確に要点を伝える共通の道具や方法論にはあまり期待がもてない。

一方、個人に対してエビデンスが正確に伝わらない場合でも、社会全般にエビデンスから得られた施策や思考パターンなどを流布させることは可能である。「説明のジレンマ」が生じにくい(と考えられる)レベルの情報量を比較的短時間で繰り返し伝達することにより、形骸的な部分は容易に伝播する。ラジオ・テレビのコマーシャルや広告チラシによる、日常生活における知識の伝播がそれである。斬新さはないが、既に一般に定着しているメディアであり、情報の受け手の認識力のトレーニングや特別なスキルは必要ない。しかしこの方法は、一つの現実的な対応策であるが、制御不能の過剰反応を生じる危険性も伴う。情報の受け手の幅が極めて広いため、エビデンスから得られた本質的な情報を正確に伝達できない可能性もあり、ときに情報内容が意図せず誇張されたり、伝達内容が多様化し全く異なる方向に認識されるという問題も生じ得る。

同じ情報を流した場合に生じるこれらの問題は、受け手側の認識上の問題であるが、その原因は多種多様であり全体としての解決方法はなさそうである。

コミュニケーションの問題は、意思を伝えようとする側とその受け手が存在する限り、常に不確定な要素が介在するため、公衆を対象とした優れたキャッチ・フレーズがいつも見つかるとは限らない。

(2) 知識伝播と社会認識

科学的な意味やその背景が伝わらなくとも、特定の機会を通じて、エビデンスの本質を社会に伝達・認識させることは可能である。在住のヘルス・ケア専門家の協力を得ることにより、地域のコミュニティーや学校教育の場を通じて、場に適した情報を効率よく提供し共有できる。直接顔を合わせた双方向性の指導は、一般を対象とした一方向性の情報よりも、大きな影響力を持つことは経験的に明らかにされている (Redelmeier 1993)。

また、日常診療の変更に関して、地域の専門家による勧告は国が公表したエビデンスに基づく一般的なガイドラインよりも強い影響力を持つことも報告されている (Lomas 1991)。このような背景を考えると、上述のマスメディアを用いた、大まかな知識伝達を、こうした地域レベルのきめ細かなフォローアップにより支援することが最も効率よく社会認識を醸成できる方法のように思われる。開発途上国における筆者らのヘルス・プロモーションの経験からも、このような国レベルのマクロなキャンペーンと地域レベルの実践的なフォローアップによるアプローチが、社会認識に有効であることは十分理解できる。

エビデンスの知識伝播と社会認識において、もうひとつ課題がある。エビデンスを効果的に伝えるコンテンツの作成がそれである。ほとんどの保健医療に関わる論文は、研究者により他の研究者のために書かれたものである。これを利用者にわかりやすく伝達する方法が必要である。すなわち、伝統的な紙媒体の、また専門用語を用いて精緻に記述された論文は、恐らくほとんどの人にとっては全くつまらない読み物であり、明確にエビデン

スが記述してあったとしてもほとんど意味をなさない。このような場合には、セールス、初等教育、グラフィックデザインや広報を専門とする人々の知識や技術を学び協力する必要がある (ハダット 2004, p.118)。特に、次世代の担い手らの特質も踏まえて、適切なメディアを利用することが肝要である。初等教育においてエビデンスについて「つたえる」ことは一見困難かのように思えるが、適切なメディアがあればその本質部分をうまく伝えることは可能と思われる。

社会認識が伴わないために、多大のリソースをつぎ込んで構築したエビデンスの宝庫が公益に資さないとすれば、これは深刻な問題である。上述のいずれの方法によるにしても、その基本として、科学の公衆理解 (public understanding of science) の振興が鍵となる。このあたりの方策と議論については、雑誌「臨床評価」の特集「ナレッジ・マネジメントと情報の共有化」に詳しい²⁹。また、集団を対象とするEBHにナレッジ・マネジメントの概念等を導入し、エビデンスに基づく意思決定、および研究から実践へのアプローチについてはミュア・グレイの著書 (2版) が参考となる (ミュア・グレイ 2005³⁴)。この意思決定や研究から実践へのアプローチの重要性と、同書で述べられている「正しいことを正しく行う (Doing the right things right)」ことの意味するところは深い。

また、エビデンスを適切に伝播し社会認識を喚起するための評価コミュニケーションは、健康問題をはじめとするあらゆる領域におけるリスク・マネジメントの視点からも重要である。さらに、EBMの方法論と問題点、現代科学論、科学哲学の視点からの考察は斉尾らの総説が参考になる³⁵。

(3) 量か質か？

RCTやSRの話をするとき必ず昔ながらの論争が噴出する。量的研究 (RCTとSRにも応用される統計学的手法の適用が可能な研究) 対、質的研究 (個人や社会の内面の定性的な側面を対象とする研究) の対峙、いわゆる研究パラダイム (対象の把握の仕方) 論争である。

RCTやエビデンスに纏わる多くの論争については、量か質かのどちらかに偏った保守的な考え方

に加え、単に、互に各々の評価研究法を未経験であることによる不理解があるように思われる。あるいは、量的研究においては数学や数理統計が共通言語と成り得るが、質的研究にはこれに対応する共通言語がないことがコミュニケーションを難しくしている原因かもしれない。

様々な議論は省略するが、明白にいえることは、RCTやSRはむしろ定性的・質的な研究を支援するものであり、あらゆる既存研究の成果を相補的に高めるものであって、決して既存研究と齟齬きたすものではないということである。必ずしも一般的ではないが、医学界に強調されてきた「根拠」や「統計」「科学性」に対する補完的な意味をもつ考えとしてNBM (Narrative-based Medicine) も提唱されている(グリーンハル2001)。質的研究法により量的研究の欠陥が明らかになることもあるので、量的研究と質的研究両者のバランスよい協調は本来の目的(対象から得られる真実を明らかにすること)への近道である。

1993年のPopeらの論文(質的研究をしている社会学者と保健医療サービス研究部門の責任者の間で交わされた架空の議論を哲学者らの問答のようにした内容)を契機に、保健医療の分野における質的研究についての注目が高まり、British Medical Journal誌の連載を経て1996年に単行本が出版された。この世界的に著名な医学誌に連載が始まってから、イギリスでは保健医療分野で質的研究が盛んに行われるようになった。質的研究が様々な問題を理解することに役立つことが注目され、2000年には改訂2版も上梓され、さらに内容を充実させている。質的研究においても研究の質の問題が生じ、質の「良いもの」と「悪いもの」を区別する方法を模索することとなった。英国NHS Research and Development Programmは、1998年頃医療技術評価に関する質的研究手法を再検討するための研究費を付けたが、残念ながら現時点でまだガイドラインは出ていないようである。

このように最近の研究上の疑問に答えるためのより一貫した効果的な方策により、定量的アプローチと定性的アプローチが併合される方向性が模索されている(Pope 1995)。2005年10月、メルボルンで開催された第13回Cochrane Colloquiumにお

いて「エビデンスは既に十分か?」との疑問のもとに議論が行われ、得られたエビデンスは単独ではなく、必要とされる場で実際に利用可能な形で研究の成果と統合されるべきであるとの再確認がなされた(Summerskill 2005)。Web検索でこの方面の動向を眺めると英国Leicester大学のQual-Quan Evidence Synthesis Groupなどがこの革新的な境界領域の研究を進めていることがわかる³⁶。ここでは、蓄積したEBHにおけるSRの経験をもとに、質的研究論文のSRを行い、質的エビデンスを併合する研究を行っている。また、2006年6月末、EBH推進の担い手となっているOxford大学のCentre for Continuing Professional Developmentは、Qualitative Research and Systematic Reviews Workshopを開催され³⁷、コクラン共同計画のCochrane Qualitative Methods Review Groupに経験のフィードバックを行うようである。EBM/EBHの系譜を踏まえ、今後の研究動向に注目したい。

おわりに

わが国における人文社会科学領域におけるエビデンスの産出とその情報基盤の整備は、次世代に向けての我々の責務である。歴史的考察から得た教訓として、政府は明確な戦略を掲げたくて、ガイドラインや法的な整備を進めるとともに、適切な予算確保の仕組みを構築する必要がある。また同時に、評価研究を行うわれわれは諸外国の評価研究者や既に活動を推進しているグループとの連携を深める必要がある。各々の役割と専門性を活かし、ひとつひとつ、社会認識を醸成するための布石を置き続けてゆくことが肝要である。澎湃たる未来は、次世代に向けた長期的視野に立つ施策と個人の科学理解および社会認識能力の高揚にゆだねられている。

注記

- 1 EBMにおけるevidenceは、「バイアスのない方法により得られたデータを、バイアスのない方法で分析して得られた結果」の総称である。介入結果のみならず診断や予後についても産出され、その産出の方

- 法や研究の質によりそのレベルは異なる。「証明力」「証拠」「科学的根拠」「科学的事実」などと訳されることがあるが現時点で定訳はない。本稿では「エビデンス」を用い、一般的用語のevidenceと区別した。
- 2 これまでに系統的な調査は知られていないが、各分野の専門家らとの情報交換によりこの状況は共通の課題として認識される。筆者らの調査からもこの状況が確認されている。
 - (1) 津谷喜一郎、津富宏、正木朋也 (2003) 「人文社会科学分野におけるランダム化比較試験などの研究デザイン使用の現状」、『厚生労働科学研究費補助金 (医療技術評価総合研究事業) 「日本におけるEBMのためのデータベース構築及び提供利用に関する調査研究 (H14-医療-067) 平成14年度 総括・分担研究報告書」 (主任研究者：丹後俊郎)』、183-189
 - (2) 正木朋也、津富宏、三砂ちづる、加藤則子 (2003) 「青少年暴力に関わる国内研究動向の文献的検討」、『厚生労働科学研究費補助金 (こころの健康科学研究事業) 「思春期における暴力行為の原因究明と対策に関する研究 平成14年度研究報告書」 (主任研究者：小林秀資)』、85-98
 - 3 本部のホームページは英語を含め7ヶ国語で閲覧可能である (<http://cochrane.org/>)。一方、国内ではJANCO (JAPANESE informal Network for COchrane Collaboration) に情報が集約されている (<http://cochrane.umin.ac.jp/>)。
 - 4 本稿では狭義のEBMには殆ど触れていない。医療現場の実践としてのEBMについては多数の書籍が出版されているので、それらを参照されたい。
 - 5 原本 J.A.Muir Gray (1996) Evidence-Based Healthcare - How to Make Health Policy and Management Decisions. Churchill Livingstone. なお、本書の日本語訳は1999年に発行されている。また、2001年に第二版が上梓され、その日本語訳も2005年に出版されている。
 - 6 5のMuir Grayの本の表現はhealthcareが一語となっている。この他、health care、あるいは、HealthCareとしている場合も見受けられる。
 - 7 ファイヨール著、山本安次郎訳 (1985) 『産業ならびに一般の管理』、ダイヤモンド社
アンリ・フェイヨール著、都筑栄訳 (1964) 『産業並に一般の管理』、風間書房
 - 8 John Wiley and Sons, Inc (Cochrane Library Customer Services Advisor Alison Robb on: +44 (0) 1243 843379, e-mail: arobb@wiley.co.uk) に直接申し込むほか、日本では、南江堂洋書部 (Tel: 03-3811-9950, Fax: 03-3811-5031, e-mail: nkdyosho@nankodo.co.jp) が代理店となっている。
 - 9 Centre for Reviews & Dissemination, University of York, UK, (<http://www.york.ac.uk/inst/crd/>)
 - 10 日米欧三極のハーモナイゼーション会議 The International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH: <http://www.ich.org/>) が作成したガイドラインのひとつ。日英対訳および関連情報は以下に詳しい。 <http://home.att.ne.jp/red/akihiro/index.htm>
 - 11 原本: Alejandro R. Jadad (1998). Randomized Controlled Trials: A User's Guide. BMJ Books.
 - 12 Journal of Negative Results in Biomedicine, <http://www.jnrbm.com/>
 - 13 International Clinical Trials Registry Platform (ICTRP) (<http://www.who.int/ictip/>)
 - 14 UMIN臨床試験登録システム (<http://www.umin.ac.jp/ctr/>)
 - 15 財団法人日本医療情報センター (<http://www.japic.or.jp/>)
 - 16 日本医師会治験促進センター (<http://www.jmacct.med.or.jp/>)
 - 17 「患者や健康な人を対象とした治療を兼ねた試験」が「臨床試験」であり、そのうち特に新薬開発 (医薬品製造承認申請) を目的とした「臨床試験」を「治験」と呼ぶ。
 - 18 臨床試験の登録・結果公開に関する実施要領 2005 <http://www.jpma.or.jp/policy/pdf/rinsyou.pdf>、同 Q&A <http://www.jpma.or.jp/policy/pdf/rinsyouqa.pdf>
 - 19 企業合併にともない一時中断したが現在も以下に引き継がれている。 (<http://ctr.gsk.co.uk/>)
 - 20 盲検 (blinding) という用語は、盲人蔑視の視点が含まれるとして日英ともに使用を控えるべきとの議論もある。代案として「遮蔽」「マスクキング」などが提案されている。
 - 21 著者らは「隠蔽」と訳している。隠蔽と訳される場合もある。
 - 22 臨床研究の現場で、二重盲検ランダム化比較試験という語は使わない。盲検化すればランダム化すること

- とは通常包含されるため、ランダム化を省略して二重盲検比較試験あるいは、比較も省略し二重盲検試験と呼ばれる。
- 23 日米欧三極のハーモナイゼーション会議 (ICH) での協議を経た「医薬品の臨床試験の実施に関する基準 (Good Clinical Practice: GCP)」や「疫学研究に関する倫理指針」などがある。いずれも、ヒトと対象とする医学研究の倫理的原則であるヘルシンキ宣言にその起源をみることができる。
- 24 クロス・オーバー法: 最も単純な例として、二種類の介入AとBの比較を行う際、ABの順に実施する群とBAの順に実施する二の群を設け、集計解析時にAとBの介入をそれぞれまとめてA群とB群を比較する方法。
- 25 Oxford Centre for Evidence-based Medicine (<http://www.cebm.net/index.asp>)
- 26 GRADE working group (<http://www.gradeworkinggroup.org/>)
- 27 原本: GRADE Working Group (2004). Grading Quality of Evidence and Strength of Recommendations. *BMJ*,328 (7454) :1490-7.
- 28 Begg,C.他 (著)、津谷喜一郎、小島千枝 (訳) (1997)「無作為化比較試験の報告の質を改善する方法 CONSORT声明」『*JAMA<日本語版>*』、74-78
- 29 Rennie,D. (著) 津谷喜一郎、小島千枝 (訳) (1997)「無作為化比較試験をいかに報告するか CONSORT声明」『*JAMA<日本語版>*』、78-79
- 30 中山健夫 (2001)「EBMを指向した「診療ガイドライン」と医学データベースに利用される「構造化抄録」作成の方法論の開発とそれらの受容性に関する研究」、『平成13年度 総括・分担研究報告書』、厚生科学研究費補助金 21世紀型医療開拓推進研究事業 (<http://mhlw-grants.niph.go.jp/>)
- 31 同上、『平成14年度 (2002) 総括・分担研究報告書』
- 32 同上、『平成15年度 (2003) 総括・分担研究報告書』
- 33 臨床評価刊行会 (2002)「特集: ナレッジ・マネジメントと情報の共有化」、『臨床評価』、29 (2・3) :222-305
- 34 原本: J.A.Muir Gray (2001). Evidence-Based Healthcare - How to Make Health Policy and Management Decisions. 2nd Ed., Churchill Livingstone.
- 35 斉尾武郎、栗原千絵子 (2001)「Evidence-based medicineの現代科学的考察 EBM questioned; a metascientific criticism」『臨床評価』、29 (1) :185-201
- 36 Qual-Quan Evidence Synthesis Group, Biostatistics & Genetic Epidemiology Group, part of the Department of Health Sciences, at the University of Leicester. <http://www.hs.le.ac.uk/research/qualquan/index.htm>
- 37 http://www.conted.ox.ac.uk/cpd/healthsciences/courses/short_courses/qs/speaker.asp

参考文献

- Feinstein,A.R. (1968). Clinical judgement. Huntington, NY: RE Krieger Pub. Co.
- Guyatt,G.H. (1991). Evidence-based medicine. ACP Journal Club March/April,A-16.
- Jones,R. (1995). Who does qualitative research? *BMJ*,311,2.
- Lomas,J., Enkin,M., Anderson,G.M., Hannah,W.J., and Vayda,E. (1991). Opinion leaders vs audit and feedback to implement practice guidelines. *JAMA*,265,2202-7.
- McNutt,R.A., Evans,A.T., Fletcher,R.H., and Fletcher,S.W. (1990). The effects of blinding on the quality of peer-review. *JAMA*,263,1371-6.
- Nakayama,T., Hirai,N., Yamazaki,S. and Naito,M. (2005). Adoption of structured abstracts by general medical journals and format for a structured abstract. *J Med Libr Assoc*,93 (2) April,237-242.
- Owen,R. (1982) Reader bias. *JAMA*,247,2533-4.
- Paul,J.R. (1938). Clinical epidemiology. *J Clin Invest*,17 (5) September,539-541.
- Pope,C. and Mays,N. (1993). Opening the black box: an encounter in the corridors of health services research. *BMJ*,306,315-8.
- Pope,C. and Mays,N. (1995). Reaching the parts that other methods cannot reach: an introduction to qualitative methods in health and health services research. *BMJ*,311,42-5.
- Redelmeier,D.A., Rozin,P., and Kahneman,D. (1993). Understanding patient' decisions: cognitive and emotional perspectives. *JAMA*,270,72-6.
- Rossi,P.H., Freeman,H.E. and Lipsay,M. (1999). Evaluation: A Systematic Approach 6th Edition.

- Thousand Oaks, CA: Sage Publication.
- Sackett, D.L., Haynes, R.B., Guyatt, G.H., et al (1991). Clinical epidemiology - A basic science for clinical medicine. Second Ed. Little, Boston/Tronto/London: Brown and Company.
- Sackett, D.L., Richardson, S., Rosenberg, W. and Haynes, R.B. (1996). Evidence-based medicine: how to practise and teach EBM. Edinburgh: Churchill Livingstone.
- Smith, R. (1988). The sickening of medical research. BMJ, 296, 1079-80.
- Summerskill, W. (2005). Cochrane Collaboration and the evolution of evidence. Lancet, 366, 1760 (Comment).
- Sykes, R. (1998). Being a modern pharmaceutical company. BMJ, 317, 1172-80. (翻訳 サイクス, リチャード (2000) 「現代の製薬企業であること それには臨床試験の情報公開プログラムを欠かすことができない」、『臨床評価』, 27(3), 509-10)
- グリーンハル, トリシャ, ハーウィッツ, ブライアン 編 / 斎藤清二, 山本和利, 岸本寛史監訳 (2001) 『ナラティブ・ベイスト・メディスン』, 金剛出版
- ミュア・グレイ, J.A. (著) 久繁哲徳 (監訳) (1999) 『根拠に基づく保健医療 健康政策と経営管理の判断決定の方法』, じほう
- ミュア・グレイ, J.A. (著) 津谷喜一郎, 高原亮治 (監訳) (2005) 『エビデンスに基づくヘルスケアヘルスポリシーとマネージメントの意思決定をどう行うか』, エルゼビアジャパン
- ハダット, アレハンドロ (著) 舟喜 光一, 津谷喜一郎 (訳) (2004) 『ランダム化比較試験 100の questionsに学ぶ』, じほう
- ポーブ, キャサリン, メイズ, ニコラス (編) 大滝純司 (監訳) (2001) 『質的研究実践ガイド』, 医学書院
- 青木仕 (2000) 「構造化抄録の基礎知識」, 中嶋宏 (監訳) 津谷喜一郎, 他 (編) 『EBMのための情報戦略. エビデンスをつくる、つたえる、つかう』, 中外医学社, 82-93
- 窪田好男 (1998) 「NPM型政策評価と政府の失敗 - 地方行革との関連で - 」, 『会計検査研究』, 18:47-62
- 佐々木亮 (著) 龍慶昭 (監修) (2003) 『政策評価トレーニングブック』, 多賀出版
- 田辺智子 (2003) 「米国90年代の行政改革」, 『レファレンス』, 30-46
- 津谷喜一郎 (1995) 「情報の質の確保」 『臨床と薬物治療』, 14(7):593-7
- 津谷喜一郎 (1996) 「コクラン共同計画とシステムティック・レビュー」, 『薬理と治療』, 25(1):11-24
- 津谷喜一郎 (1999) 「EBMとコクラン共同計画」, 矢野栄二 (編) 『医療と保健における評価 Toward Evidence Based Medicine』, 南江堂, 195-217
- 津谷喜一郎 (2000a) 「コクラン共同計画とは 日本的展開へ向けて」, 『公衆衛生研究』, 49(4):313-9
- 津谷喜一郎 (2000b) 「コクラン共同計画とシステムティック・レビュー EBMにおける位置付け」, 厚生省健康政策局研究開発振興課医療技術情報室 (監修) 『わかりやすいEBM講座』 厚生科学研究所, 120-35
- 津谷喜一郎 (2000c) 「医学研究デザインの基礎」, 中嶋宏 (監修) 津谷喜一郎, 他 (編) 『EBMのための情報戦略. エビデンスをつくる、つたえる、つかう』, 中外医学社, 22-47
- 津谷喜一郎, 中山健夫, 島村治子 (訳) (2005) 「エビデンスの質とお勧め度のグレーディング」 『薬理と治療』, 33(12):1241-54
- 林謙治 (2000) 「根拠に基づく健康政策へのアプローチ」, 『公衆衛生研究』, 49(4):346-53
- 龍慶昭, 佐々木亮 (著) (2000) 『「政策評価」の理論と技法』, 多賀出版

(2006.2.28受理)

The Genealogy and Direction of Evidence-based Medicine (EBM): The Future and Role of the Cochrane Collaboration to Achieve Health Care Evaluation

Tomoya Masaki

Health Service Management,
International University of Health and Welfare Graduate School
masakit-ky@umin.ac.jp

Kiichiro Tsutani

Graduate School of Pharmaceutical Sciences, The University of Tokyo
tsutani-ky@umin.ac.jp

Abstract

Although there is a social demand for scientific evidence of public investment, experience in areas other than healthcare are quite rare in Japan. However, in the United Kingdom and the United States of America, several activities of evidence-based practice in decision making for public policy are in existence, and these evidences are steadily increasing in number. These examples are: the Campbell Collaboration, What Works Clearinghouse of US Department of Education, and the Poverty Action Laboratory in the area of official development assistance. Taking into account such a background, an overview of the genealogy of the Evidence-based Medicine (EBM) and the Cochrane Collaboration as the precedent experience is reviewed in this paper. In addition, the role and the future direction of EBM with regard to the importance of accumulating quality information as a social infrastructure and social epistemology are discussed.

Keywords

EBM, evidence, randomized controlled trial, systematic review, social epistemology

【研究ノート：依頼原稿】

刑事司法におけるエビデンスの活用：非行の予防・処遇を中心として

津富 宏

静岡県立大学

tsutomi@u-shizuoka-ken.ac.jp

要 約

刑事司法分野、とりわけ、非行の予防・処遇については、RCT（ランダム化比較試験）や擬似実験、それらに基づくSR（系統的レビュー）が行われ、エビデンスが蓄積されてきた。そこで課題となるのが、蓄積されたエビデンスをいかに活用するかである。本稿では、アメリカ合衆国において、非行の予防・処遇に係るエビデンスの活用のために提示されている2つの手法 - 〔手法1：特定のプログラムについてRCTやSRによって効果検証を行って有効性を判定し、無効な（あるいは有害な）プログラムを実務から退却させる一方、有効なプログラムの定着を図るという手法〕及び〔手法2：非行予防・処遇全般に関するSRを行い、「有効なプログラム」に関する原則を抽出し、この原則を踏まえたサービスを提供するという手法〕 - について概観しつつ、エビデンスをわが国のコンテキストに適用するという観点からいずれの手法が適しているかを検討した。その結果、手法2については、プログラムの頑健さを要求しない、わが国の実務を否定するのではなく改善する形で利用できるという利点があり、よって、わが国においてはエビデンスの活用を行うには、手法2によることが現実的であると結論する。

キーワード

エビデンスの活用、刑事司法、非行、RCT（ランダム化比較試験）、SR（系統的レビュー）

1. はじめに

刑事司法分野、とりわけ、非行の予防・処遇については、RCT（randomized controlled trial：ランダム化比較試験）や擬似実験、それらに基づくSR（systematic review：系統的レビュー）が行われ、エビデンスが蓄積されてきた¹。そこで課題となってきたのは、「エビデンスをいかに実務に活用するか」である。本稿では、この問いへの解答を求めるための模索が続く、アメリカ合衆国の

状況を紹介し、わが国への意義を検討したい。

非行の予防と処遇に関するエビデンスを活用する手法には二つある。一つは、特定のプログラムについてRCTやSRによって効果検証を行って有効性を判定し、無効な（あるいは有害な）プログラムを実務から退却させる一方、有効なプログラムの定着を図るという手法である。もう一つは、特定のプログラムについてRCTやSRを行うのではなく、非行予防・処遇全般に関するSRを行い、「有効なプログラム」に関する原則を抽出し、こ

の原則を踏まえたサービスを提供するという手法である。

アメリカ合衆国を見ると、前者の手法に基づいたエビデンス活用は急速に進展している一方、後者の手法は提案されてはいるものの実務上の応用はこれからという状況である。しかし、筆者は、欧米において蓄積されたエビデンスをわが国のコンテキストに適用するという観点からは、むしろ、後者の手法が有効であると考えられる。以下、この二つの手法について概観する。

2. 手法1：特定のプログラムの放棄・普及によるエビデンス活用

国際的な取組みであるキャンベル共同計画²を始めとして、英語を媒体とし、非行の予防・処遇に関するエビデンスを提供しているウェブサイトは少なくない。筆者が青少年暴力に関するエビデンスを提供しているウェブサイトをレビューした際には、23のウェブサイトが見出された（津富2005）。

この種のウェブサイトを代表するウェブサイトとして、司法省の少年司法非行防止局（OJJDP: Office of Juvenile Justice and Delinquency Prevention）などの支援を受け、コロラド大学暴力予防研究センター（CSPV: Center for the Study and Prevention of Violence）が運営する、Blueprints for Violence Prevention（以下、Blueprintsという。）³がある。

Blueprints for Violence Preventionは、有効な暴力防止プログラムをModelプログラムとして同定し、かつ、これらModelプログラムを実施しようという主体に研修と技術的支援を行うことを目的としている。Modelプログラムとして現在認定されているのは、強力な研究デザイン（RCTないしマッチング統制を用いた擬似実験）を用いている、暴力、非行、薬物乱用のいずれかをアウトカムとする、最低1年間にわたり効果が持続する、最低一か所において追試されるという条件を満たした、11件のプログラムである。アメリカにおいては、処遇プログラムは開発者によって登録商標化され、その利用は有償であることが多い

が、これら11件はいずれもこうしたプログラムである⁴。

こうしたウェブサイトが情報を提供したからといって、直ちに実務が影響を受けるわけではない。そこで、どのようなメカニズムによって、エビデンスが実務に反映されるかを、無効な（あるいは有害な）プログラムの放棄、有効なプログラムの定着の順に見ていこう。

（1）無効な（あるいは有害な）プログラムの放棄

刑事司法に関するプログラムは信念に基づいていることが多く、否定的なエビデンスが提供された⁵プログラム⁶であっても、放棄されることは滅多にない。しかし、否定的なエビデンスが活用され、プログラムが放棄される事例もある。そのもっとも代表的な事例が、学校における薬物防止教育の中核であるD.A.R.E. (Drug Abuse Resistance Education)である。D.A.R.E.は、1980年代初めにロサンゼルスで始まった、制服の警察官が学校に来て小学校5年・6年生を対象に、毎週1コマ、1学期間、薬物教育を行うというプログラムである。1999年の時点では、全米の学区のおよそ80%が利用していた（Petrosino 2005, Figure 1.）。

D.A.R.E.の効果検証は、1990年代に始まり、メタ・アナリシス（Ennett et al. 1994）、一次研究（たとえば、Rosenbaum et al. 1995）のいずれもが、その効果に疑問を呈した。しかしながら、こうしたエビデンスが直ちにD.A.R.E.の隆盛に影響することはなく、むしろ、1990年代を通じてD.A.R.E.を採用する学校は増え続けた。

この動きにストップがかかったのは、連邦教育省のOffice of Safe and Drug-Free Schools が2001年に発表した、List of Exemplary or Promising Safe, Disciplined, and Drug-Free School Programs⁷の影響である。このリストは、9件のExemplaryプログラムと33件のPromisingプログラムを掲げたが、D.A.R.E.は含まなかった。このリストがD.A.R.E.の利用に影響を与えたのは、「研究成果に基づいたプログラムを実施する」か、「研究成果に基づかないプログラムを実施する場合には2年以内に有効性を示す」かのいずれかを選択しないと、プログラムに対する連邦予算の配賦を縮小ないし中

止するという「誘導」と併せて提示されたからである。プログラムの効果検証を自力で行う学区はなく、D.A.R.E.が有効であると「信じる」学区も次々とD.A.R.E.から撤退し、他のプログラムへと転換した。つまり、この転換は、各学区が産出した（あるいは少なくとも理解した）エビデンスに基づいて行われたわけではない。各学区は単に、リストから受身的にプログラムを選んだのである。つまり、より上位の行政レベルの誘導によるエビデンスに基づく行政は、ローカルなエビデンス理解を経由せず、シンプルな「オン」と「オフ」のメカニズムを提供した（Weiss et al. 2005）。

D.A.R.E.に関するエビデンスが実務に影響を与えた第二のルートは、メディアを通じてである（Petrosino 2005）。D.A.R.E.が無効であることを人々が知ったのは、New Republic, Rolling Stone, New York Times, Boston Globe, U.S. News and World Report, Newsweekといった活字メディア、Dateline, 20/20といった放送メディアが、D.A.R.E.が無効であることや、ソルトレイクシティやミネアポリスといった大都市がD.A.R.E.を放棄したことを報道したことによる。D.A.R.E.を自発的に放棄し他のプログラムにスイッチした学区の実務家は、いずれもメディアから『研究』に基づくエビデンスを入手していた（Weiss et al. 2005）。

D.A.R.E.の事例は、プログラムの放棄に当たりエビデンスが活用されるメカニズムにおいては、エビデンスのローカルな産出や理解が鍵となるわけではなく、むしろ、より上位の行政レベルからの「誘導」やメディアの影響が重要であることを示唆している。

（2）有効なプログラムの普及

プログラムの放棄とプログラムの普及は表裏一体の関係にあり、行政的誘導が有効であることに変わりはない⁸。しかし、放棄と普及は2点において異なっている。第1点は、普及は、放棄と異なり、新たな予算措置を必要とすること、第2点は、プログラムが普及すべきコンテキストにおいても有効であるかどうかは、そのコンテキストにおける実施条件にかかっていることである。いずれの

点においても、先駆的な取組みを行っているのは、ワシントン州議会の外部研究機関である、ワシントン州公共政策研究所（WSIPP: Washington State Institute for Public Policy）⁹である。WSIPPは、刑事司法分野における研究を通じて、有効なプログラムの普及を図ってきた。

第1点であるが、WSIPPは、犯罪減少プログラムの費用便益分析¹⁰（Aos et. al 2001¹¹; Aos et al. 2004¹²）を行い、これをもとに議会・行政を説得するノウハウを積み上げてきた。WSIPPの研究者Aos（2005）は、政府はサービスの購入者であるから、研究者はエビデンス（＝商品）の生産ではなく消費に目を向けなければならないと主張し、商品を買ってもらうためには、次の3点が重要だと指摘している。

売り込んでくれる人。私心のない、鍵となる議員を一人か二人確保する。政治では、何を伝えるかより誰が伝えるかのほうが大事である。

売り込むタイミング。メディアの見出し記事とタイミングを合わせる。

セールストーク。経済的な観点（費用と便益）をきちんと示し、プログラムを導入する提案は無効なプログラムをカットする提案とセットとするのがよい。

このように、WSIPPの真骨頂は、ローカル・ポリティックスをいかに説得するかというノウハウにある。で強調される経済的な観点を担うのが、費用便益分析である。Aos et al.（2004）は、61件の非行の予防・処遇プログラムの費用便益分析を行い、うち38件について便益が費用を上回ることを見出し、十分に納税者の期待に応えうるプログラムがある¹³と結論付け、州議会及び州知事に対し、次の2点を提言している。

行政的誘導：研究によって有効であることが示されているプログラムに予算を配賦し、同時に、研究によって無効であることが示されているプログラムを廃止すること（そして、そのプログラムに充てられていた予算を有効なプログラムに振り替えること）。また、地方の行政体（市、郡など）が、研究に裏付けられたプログラムを優先的に実施するようなインセンティブを設けること

エビデンスの実践： 研究によって有効であることが示されているプログラムであっても、実施上の十分な品質管理がなければ、現場で有効であるとは限らないこと

が、先ほど指摘した、第2点に当たる。WSIPPIは、Blueprintsで取り上げられたFunctional Family Therapy (FFT) などのプログラムを州規模で追試を行っているが、その結果、Barnoski (2004) は、職務能力を備えた担当者が実施するか否かで、プログラムの効果が異なることを見出した。たとえば、FFTは、職務能力を備えた担当者が実施した場合、重罪再犯を38%減少させたが、備えていない担当者が実施した場合、16%増加させた。費用便益分析に換算すると、前者は1ドルの投資あたり10.69ドルの利益、後者は4.18ドルの損失であった。WSIPPではこの研究の予備的知見 (Barnoski 2002a; 2002b) に基づき、処遇の実施の質、及び担当者の質を担保し、再犯率・プログラム修了率・再犯リスク要因を測定することによりプロセス管理を行うための基準 (Barnoski et al. 2003) を策定した。

プログラム実施の質が重要だというこの知見は、Petrosino and Soydan (2005) によっても確認されている。彼らは、1945年から1993年にかけて公表された、個人を介入単位とし、犯罪行動をアウトカムとする、RCTないし擬似RCT (交替での割付) を300件見出し、(19件を情報不足として除外して) 281件をサンプルとするメタ・アナリシスを行った。その結果、サンプルすべてを統合した効果値は「.10」であるが、プログラム開発者自身が評価者である場合には効果値が「.47」と大きいことを見出した。この結果は、処遇がモデルどおりに行われることの重要性、すなわち、処遇の忠実さ (fidelity) の重要性を示していると解釈できる^{14, 15}。

以上、プログラムの普及に当たりエビデンスが活用されるメカニズムにおいては、(政策決定者に対しては) 費用便益分析による経済的な根拠をもった説得、(実務家に対しては) プロセス管理による処遇の忠実度の確保が鍵となることが明らかとなった。

さて、処遇の忠実さがこれほどまでに処遇効果に大きな差を及ぼすとすると、従来、Best Practiceといわれてきたプログラムの、実施コンテキストを超えた頑健性¹⁶に疑問を呈さざるを得ない。そこで、有用なのが手法2によるエビデンスの活用である。

3. 手法2：「有効なプログラム」に関する原則の適用によるエビデンスの活用

特定のプログラムについてRCTやSRを行うのではなく、非行の予防・処遇全般についてSR、とりわけ、メタ回帰分析を行うことで、一般的にどのようなプログラムが有効かという原則を抽出することができる。Howell and Lipsey (2004) は、これまで実施されたSR (例えば、Lipsey 1992; Lipsey and Wilson 1998) に従い、プログラムの有効さは、主たるサービス、従たるサービス、サービスの密度・期間など実施上の要素、対象者とのマッチングの4つの要素によって決まり、これらの要素について有効な内容を提供できれば、そのプログラムの再犯率を最小化することができると予測する。例えば、(非行予防プログラムとして) 親の訓練・カウンセリング、対人関係訓練、個人別学習支援、(審判命令として) 家族カウンセリング、個人別学習支援、メンタリング、行動マネジメント、ライフ・スキル、密接な監督、認知行動療法、15週間以上にわたって31日間以上のプログラム提供、ハイリスクの年少の少年といった内容である (Lipsey 2003)。HowellとLipseyは、こうした要素を組み合わせることで、プログラムの有効さを評定するためのツールである、Standardized Program Evaluation Protocol (SPEP) を提案している。表1に、SPEPに基づき、主たるサービスを家族カウンセリングとした場合の、評定用フォームを示した。

このフォームは、プログラムの設計図としても、また、自らの実施しているプログラムの採点表としても用いることができる。SPEPの得点は、再犯率の予測と表裏一体の関係になっている。つまり、得点が高がるほど、予測される再犯率が下がる仕掛けとなっている。つまり、SPEPに基づい

表1 「有効なプログラム」の原則に基づくプログラムの有効さの評定

審判命令によるプログラム		60点
主たるサービス： 家族療法		
従たるサービス 親の訓練 10点 個人カウンセリング 2点 なし 0点	薬物/アルコールカウンセリング 6点 メンタリング 4点	<input type="checkbox"/>
サービスの期間 なし 0点 67% 6点	33% 3点 100% 9点	<input type="checkbox"/>
直接的なサービス提供の日数 なし 0点 67% 8点	33% 4点 100% 12点	<input type="checkbox"/>
過半数の少年のリスクレベル 低リスク 2点 高リスク 4点		<input type="checkbox"/>
少年の年齢 14歳以下 5点 15歳以上 2点		<input type="checkbox"/>
合計点		<input type="checkbox"/>

(出所) Howell and Lipsey (2004) Figure 2

表2 「有効なプログラム」の原則に基づく再犯率の予測

審判命令によるプログラム		再犯率
条件		
プログラムに参加していない少年の再犯率(統制群の追跡6ヵ月後の値)		.40
平均的なプログラムに参加している少年の再犯率		.34
平均以上の主たるサービスを受けた少年の再犯率		.32
平均以上の主たるサービスと最良の従たるサービスを受けた少年の再犯率		.28
平均以上の主たるサービスと最良の従たるサービスを、最良の実施条件(実施期間及び接触日数)で受けた少年の再犯率		.24
平均以上の主たるサービスと最良の従たるサービスを、最良の実施条件(実施期間及び接触日数)で、対象者適性(リスク及び年齢)がマッチした少年が受けたときの再犯率		.21

(出所) Howell and Lipsey (2004) Table 1

てプログラムを設計すれば、より低い再犯率が予測できる。手法2に従えば、特定のBest Practiceに対する厳密な忠実さを求められることはない。むしろ、大きな意味での原則に従っている限りにお

いて、非行の予防・処遇プログラムは、一定の再犯率の減少をもたらすことが統計的に予測できるのである。

表2は、 から の要素を適切に組み合わせるこ

とによって予測される、再犯率がどんどん下がっていくことを示している。つまり、表2は、表1において得点の高いプログラムを設計すればするほど、予測される再犯率が低くなることを示したものであるといえる。

Best Practiceといわれるプログラムも、ひとつだけの要素で出来上がっているわけではなく、複数の要素の組み合わせである。手法2は、プログラムの有効さの「素」^{もと}を四つの要素にいったん分解して組み立てなおそうという発想であり、予防・処遇プログラム提供者に、自らが提供しているプログラムがどのように有効である(ない)かを本質的に理解させるという機能をもっている。HowellとLipseyは、SPEPに基づく少年司法サービスの改善を、ノースカロライナ州で展開中である¹⁷。

4. おわりに

海外で産出されたエビデンスを、わが国のコンテキストで活用しようというとき、いずれの手法がより有用であろうか。手法1は、Best Practiceの細部にいたるまでの再現を求める。現時点までの研究成果を見ると、プログラムは実施コンテキストに対して頑健ではなく、その効果は、実施の忠実さに大きく依存している。海外で開発されたプログラムを、コンテキストの大きく異なるわが国において忠実に実施し、効果を挙げることは可能なのだろうか。しかも、Best Practiceはいずれも有償のパッケージプログラムであるため、これをわが国に導入するためには、新たな予算措置を講じなければならない、費用負担の問題が大きい。

一方、手法2は、SRで見出された一般的な原則に基づいて、サービスを構築するという考えに立っているため、手法1のようにプログラムの頑健さを要求しない。また、SPEPは、現行のプログラムの評定を出発点にしてその改善を図る仕組みであって、新たなプログラムの導入を要求するわけではないから、直ちにわが国の実務を直ちに否定しない。

以上、二つの手法を比較すると、わが国においては、手法2に立ってエビデンス活用を図ること

が現実的であると考えられる。

女児の殺人など、悲惨な事件が大きく報道されるたびに、わが国のメディアでは、海外の犯罪対策の安易な「導入」が提唱されがちである。現時点において、RCTやSRの実践を欠いているわが国は、エビデンスの「輸入国」として、エビデンス産出よりも、むしろ、エビデンス活用の手法についての議論を深める必要がある。

注記

- 1 たとえば、最低50単位が実験群あるいは統制群に割り付けられた、犯罪をアウトカムとしている、英語で発表されたRCTは、1957年から1981年の間に35件、1982年から2004年の間に83件あることが報告されている（Farrington 1983; Farrington and Welsh 2005）。また、Petrosino and Soydan（2005）は、犯罪行動をアウトカムとする、メタ・アナリシスが50件あることを報告している。
- 2 現時点で、完了し公表されたSRが2件（Scared Straight 及びブートキャンプ）、完了し審査中のSRが12件、プロトコルが承認されたSRが7件、プロトコルが審査中のSRが11件、タイトルのみ承認されたSRが9件である（CCJG Steering Group 2005）。ブートキャンプとは、矯正施設を、新兵訓練所に模して運営することにより、更生を図るという処遇である。
- 3 <http://www.colorado.edu/cspv/blueprints/>
- 4 処遇プログラムを登録商標化する流れは、心理療法の商業化に淵源を持つ。
- 5 社会政策に関するエビデンス提供するウェブサイトには、エビデンス提供の仕方を見ると、2種類ある。一つは、効果検証の結果を、プログラムが有効であるか無効であるかにかかわらず報告するウェブサイト、もう一つは、有効なプログラムのみをリスト化し、推奨するウェブサイトである。前者は、無効な（あるいは有害な）プログラムを明示的に示すことにより、後者は、無効な（あるいは有害な）プログラムを推奨リストから外すことにより、無効な（あるいは有害な）プログラムを実務から排除することを狙っている。
- 6 たとえば、Scared Straight（エビデンスは、Petrosino et al. 2003）は有害、ブートキャンプ（エビデンスは、

Wilson et al. 2005) は無効と、否定的エビデンスが示されたにもかかわらず放棄される気配がない。

7 <http://www.ed.gov/admins/lead/safety/exemplary01/exemplary01.pdf>

8 先述した、教育省と同様の取組みは、司法省においても行われている。同省のOJJDPはすべてのプログラムにロジック・モデルを導入し、アウトプットの業績指標としてBest Practiceモデルを用いているかどうかを含めている。

Best Practiceは研究及び評価によって少年非行を防止・抑止する効果があると認められた対策及びプログラム。Best Practiceモデルは狙いとするアウトカムを達成することが、厳密な評価及び追試によって示されたプログラム・モデルを含む。モデルプログラムについての情報は、OJJDPのModel Programs Guide、CSPVのBlueprints、健康福祉省薬物乱用・精神衛生管理局 (SAMHSA: Substance Abuse and Mental Health Service Administration) のModel Programs、各州のモデルプログラムに関する情報源などから得られる」と定義されている。ロジックモデルは、OJJDPのすべての助成プログラムを対象とするため、大きな影響を与えるものと思われる。

9 <http://www.wsipp.wa.gov/>

10 刑事司法における費用便益分析において困難なのは、便益の算出 (犯罪のもたらす費用の算出) である。この分野の研究は、Cohenが研究を集大成 (Cohen 2005) したほか、医療経済学を応用した論考 (例えば、Dolan et al. 2005) も現れ、活況を呈している。2002年には、コクラン共同計画とキャンベル共同計画のジョイントの経済学方法論グループ (Campbell & Cochrane Economics Methods Group <http://www.c-cemg.org/>) が結成されている。

11 <http://www.wsipp.wa.gov/rptfiles/costbenefit.pdf>

12 <http://www.wsipp.wa.gov/rptfiles/04-07-3901.pdf> (本文) <http://www.wsipp.wa.gov/rptfiles/04-07-3901a.pdf> (補遺A) <http://www.wsipp.wa.gov/rptfiles/04-07-3901b.pdf> (補遺B)

13 例えば、最も有効とされているのは、Dialectical Behavior Therapyという認知行動療法であり、少年一人当たり、便益が32807ドル、費用が843ドルで、差し引き31243ドルの利益があると試算されている。

14 もちろん、評価者自身の利害の意識的・無意識的な

利害の反映の結果としても解釈することができる。

15 CSPVはModelプログラムの普及を目指し、Blueprints Conference 2006 (<http://www.blueprintsconference.com/>) を2006年5月に開催する。このカンファレンスでもプログラムを忠実に実施するための工夫が一つの中核的なテーマである。

16 例えば、Curtis et al. (2004) の、Multisystemic Therapyに関するメタ・アナリシスでは、プログラム開発者が担当者の研修とモニタリングに主体的に取り組む、かつ、担当者が博士課程の学生だった場合と、プログラム開発者が研修とモニタリングに主体的に取り組まず、かつ、担当者が修士課程の学生だった場合を比較して、前者では効果値が「.81」、後者では「.28」と大きな差があることを報告しているが、この結果も、忠実さの違いを反映しているように思われる。最近発表された、Multisystemic Therapyのメタ・アナリシス (Littell et al. 2005) において効果が見出されなかったのも、プログラムが普及するにつれ、忠実度が担保できなくなった結果だと解釈することもできる。

17 <http://www.ncdjjdp.org/jcpc/evaluation.html>

参考文献

Aos, S. (2005). Getting what you pay for with evidence-based criminal justice: The role of the state and local government purchaser. Presented at the 57th meetings of the American Society of Criminology.

Aos, S., Lieb, R., Mayfield, J., Miller, M., and Pennucci, A. (2004). *Benefits and Costs of Prevention and Early Intervention Programs for Youth*. Washington State Institute for Public Policy. 2006年2月12日に<http://www.wsipp.wa.gov/pub.asp?docid=04-07-3901>から入手

Aos, S., Phipps, P., Barnoski, R., and Lieb, R.. (2001). *The Comparative Costs and Benefits of Programs to Reduce Crime, v 4.0*. Washington State Institute for Public Policy. 2006年2月12日に<http://www.wsipp.wa.gov/pub.asp?docid=01-05-1201>から入手

Barnoski, R. (2002a). *Washington State's Implementation of Aggression Replacement Therapy: Preliminary Findings*. Washington State Institute for Public Policy. 2006年2月12日に<http://www.wsipp.wa.gov/pub.asp?docid=02-06-1201>から入手

- Barnoski, R. (2002b). *Washington State's Implementation of Functional Family Therapy for Juvenile Offenders: Preliminary Findings*. Washington State Institute for Public Policy. 2006年2月12日に<http://www.wsipp.wa.gov/pub.asp?docid=02-08-1201>から入手
- Barnoski, R. (2004). *Outcome Evaluation of Washington State's Research-Based Programs for Juvenile Offenders*. Washington State Institute for Public Policy. 2006年2月12日に<http://www.wsipp.wa.gov/pub.asp?docid=04-01-1201>から入手
- Barnoski, R., Aos, S., and Lieb R. 2003. *Recommended quality control standards: Washington State research-based juvenile offender programs*. Washington State Institute for Public Policy. 2006年2月12日に <http://www.wsipp.wa.gov/pub.asp?docid=03-12-1203>から入手
- CCJG Steering Group. (2005). Progress report for Toronto CCJG Steering Group meeting, November 2005. Memorandum.
- Cohen, M. A. (2005). *The Costs of Crime and Justice*. London: Routledge.
- Curtis, N. M., Ronan, K. R., and Borduin, C. M. (2004). Multisystemic treatment: A meta-analysis of alternative education programs. *Crime and Delinquency*, 41, 219-234.
- Dolan, P., Loomes, G., Peasgood, T. and Tsuchiya, A. (2005). Estimating the intangible victim costs of violent crime. *British Journal of Criminology*, 45, 958-976.
- Ennett, S. T., Tobler, N. S., Ringwalt, C. L., and Flewelling, R. (1994). How effective is drug abuse resistance education? A meta-analysis of Project DARE outcome evaluations. *American Journal of Public Health*, 84 (9), 1394-1401.
- Farrington, D. P. (1983). Randomized experiments on crime and justice. In M. Tonry and N. Morris (Eds.), *Crime and Justice, Vol. 4*. Chicago: University of Chicago Press, 257-308.
- Farrington, D. P., and Welsh, B. C. (2005). Randomized experiments in criminology: What have we learned in the last two decades. *Journal of Experimental Criminology*, 1, 9-38.
- Howell, J. C., and Lipsey, M. W. (2004). A practical approach to evaluating and improving juvenile justice programs. *Juvenile and Family Court Journal*, Winter 2004, 35-48.
- Littell J. H., Popa M., and Forsythe, B. (2005). *Multisystemic Therapy for Social, Emotional, and Behavioral Problems in Youth Aged 10-17*. Campbell Collaboration. 2006年 2月 13日 に http://www.campbellcollaboration.org/doc-pdf/Littell_multhe_rev.pdf から入手
- Lipsey, M. W. (1992). Juvenile delinquency treatment: A meta-analytic inquiry into the variability of effects. In T. D. Cook, H. Cooper, D. S. Cordray, H. Hartman, L. V. Hedges, R. J. Light, et al. (Eds.), *Meta-Analysis for Explanation*. New York: Russell Sage Foundation, 83-127.
- Lipsey, M. W. (2003). Standardized Program Evaluation Protocol development: The evidence base. 2005年12月29日に、<http://www.ncdjdp.org/news/ppt/HowellAreaConferencePresentation.ppt> から入手
- Lipsey, M. W. and Wilson, D. B. (1998). Effective interventions with serious juvenile offenders: A synthesis of research. In R. Loeber and D. P. Farrington (Eds.), *Serious and violent juvenile offenders: Risk factors and successful interventions*. Thousand Oaks, CA: Sage Publications, 313-345.
- Petrosino, A. (2005). 「D.A.R.E. and scientific evidence: a 20 years history」, 『犯罪社会学研究』, 30: 72-88
- Petrosino, A. and Soydan, H. (2005). The impact of program developers as evaluators on criminal recidivism: Results from meta-analyses of experimental and quasi-experimental research. *Journal of Experimental Criminology*, 1, 435-450.
- Petrosino, A., Turpin-Petrosino, C. and Buehler, J. (2003). "Scared Straight" and Other Juvenile Awareness Programs for Preventing Delinquency. Campbell Collaboration Review. 2006年2月13日に<http://www.campbellcollaboration.org/doc-pdf/ssrupdt.pdf> から入手
- Rosenbaum, D. P., Flewelling, R. L., Bailey, C. L., and Wilkinson, D. L. (1994). Cops in the classroom: A longitudinal evaluation of Drug Abuse Resistance Education (D.A.R.E.). *Journal of Research in Crime and Delinquency*, 31, 3-31.
- 津富宏 (2005) 「実務家が、青少年の暴力予防プログラムの高価に関するエビデンスを入手する - ウェブを通じたレビュー成果の提供 - 」, 『保健医療科学』,

54 (2) : 127-134.

Weiss, C. H., Murphy-Graham, E., and Birkeland, S. (2005). An alternate route to policy influence: how evaluations affect D.A.R.E. *American Journal of Evaluation*, 26, 12-30.

Wilson, D. B., MacKenzie, D. L., Mitchell, F, N. (2005).

Effects of Correctional Boot Camps on Offending. Campbell Collaboration. 2006年2月12日にhttp://www.campbellcollaboration.org/doc-pdf/Wilson_bootcamps_rev.pdf から入手

(2006.2.28受理)

Utilization of Evidence in Criminal Justice with a Focus on Prevention and Treatment of Delinquency

Hiroshi Tsutomi

University of Shizuoka
tsutomi@u-shizuoka-ken.ac.jp

Abstract

Scientific evidence produced by RCTs (randomized controlled trials) and quasi-experiments and SRs (systematic reviews) has been accumulated in criminal justice, especially as regards delinquency prevention and treatment. We now face the challenge of how to utilize the evidence so that it can be put into practice. In this article, the author describes two approaches to evidence utilization so far developed in the U.S. and discusses the relative applicability of these approaches to a Japanese context. The first approach is a model program approach which encourages practitioners to discard ineffective and harmful programs and to adopt effective model programs based on the results of RCTs and/or SRs while the second approach is a model principle approach which encourages practitioners to assess and improve their programs in terms of effective program principles derived from SRs. The second approach is preferred as a method of utilizing evidence in Japan firstly because it does not require the robustness of programs across contexts and secondly because it allows us to improve Japanese criminal justice practices without replacing them.

Keywords

utilization of evidence, criminal justice, delinquency,
RCT (randomized controlled trial), SR (Systematic Review)

【研究ノート：依頼原稿】

エビデンスに基づく教育 アメリカの教育改革とWhat Works Clearinghouseの動向

田辺 智子

国立国会図書館

s-tanabe@ndl.go.jp

要 約

アメリカの「子どもを一人も落ちこぼれにしないための2001年法」(NCLB法)は、教育現場にエビデンスに基づく施策の採用を求めている。エビデンスとして特に重視されるのが、実験デザインを用いた評価研究である。教育省は、評価研究に関する信頼できる情報を提供することを目的に、What Works Clearinghouse (WWC) というウェブサイトを開設した。評価研究の系統的レビューを行い、その結果をエビデンス・レポートとして掲載している。WWCはまだ構築の途上であるが、今後、どのようにコンテンツが充実され、教育現場に採用されていくのか、動向が注目される。

日本の教育界にとっても、エビデンスを重視するNCLB法の発想は示唆的である。わが国でも、信頼性の高い評価研究を蓄積し、系統的レビューを行っていく意義は大きい。

キーワード

エビデンス、プログラム評価、実験デザイン、教育、What Works Clearinghouse

はじめに

エビデンスに基づく施策の採用を目指す最近の評価研究の流れにあって、新しい動きとして注目されるのが、アメリカの教育改革とWhat Works Clearinghouse (WWC) の動向である。ブッシュ政権は、州や学校現場に対し、エビデンスに基づいた意思決定を求める法改正を行っており、そうした意思決定を支援するために、評価研究の系統的レビュー (Systematic Review) を提供する情報サイトであるWWCが開設された¹⁾。

エビデンスの系統的レビューの試みとしては、保健医療分野のコクラン共同計画や、社会科学の

キャンベル共同計画が先行している。WWCも両計画に大きな影響を受けているが、コクランやキャンベルとの重要な違いは、WWCは開設の時点ですでに、コンテンツに対する需要が用意されているという点にある。WWCによって今後、充実が図られていくコンテンツは、アメリカの教育現場の意思決定に、ダイレクトな影響を与えることが予想される。

本稿では、アメリカの教育改革におけるエビデンスの重視とWWCの現状について、関連する議論の動向を含めながら紹介する。

1. アメリカの教育改革とエビデンスの重視

(1) 「子どもを一人も落ちこぼれにしないための2001年法」

2001年に誕生した現ブッシュ政権は、教育を最優先の政策課題と位置づけ、政権発足後いち早く教育改革に着手した。2002年1月には、超党派の支持を受け、「子どもを一人も落ちこぼれにしないための2001年法」(No Child Left Behind Act of 2001: NCLB法²⁾)が成立した。NCLB法の基本的目的は、アメリカの教育水準を引き上げることに置かれている。そのために、州や学校の裁量を拡大し、代わりに子どもの学力向上という「結果」に対する責任を厳しく問うアカウンタビリティ・システムが導入された。その他、具体的な施策として盛り込まれたのは、読解力の向上、親への情報と選択権の付与、教師の質の向上、学校の安全確保、テクノロジーを活用した教育の促進などである(松山2001)。

以上に加え、NCLB法は、それまでにない重要な特徴を持っていた。科学研究によって有効性が裏付けられた施策や指導法の重視である。NCLB法の条文には、100回以上も「科学的根拠のある研究」(Scientifically-Based Research)という表現が登場する。たとえば、「科学研究に基づいて総合的な読解プログラムを構築した州政府に対し予算を与える」、「不利な条件にある生徒への教育サービスの計画にあたっては、科学的根拠のある研究結果を考慮しなければならない」、「若者の暴力やドラッグを防止すると科学的に示された施策に予算を振り向ける」といった具合である(Reyna 2005; U.S. Department of Education 2002)。

NCLB法のベースにあるのは、連邦政府はこれまで教育に多額の予算を投入してきたにもかかわらず、学力向上という成果につながっていないという問題意識であった。そして、成果を上げるための手段の一つと位置づけられたのが、科学的に有効性が確認された施策の採用、つまりエビデンスの重視である。

(2) 何がエビデンスとされるのか

では、NCLB法のいう「科学的根拠のある研究」とは、具体的にどのようなものを指すのだろうか。同法の第9101条では、その定義を次のように与えている。「教育活動や施策に関する妥当で信頼できる知識を得るための、厳格・体系的かつ客観的な手続きの適用を含む研究」であって、それには、観察や実験に基づく体系的・実証的研究、仮説の検証を伴うデータ分析、実験デザイン・擬似実験デザインを用いた研究などが含まれる。

この定義からは、実験デザインや擬似実験デザインが「科学的根拠のある研究」に含まれることは読み取れるが、定義を満たす研究の要件は必ずしも明らかではない。また、一般の教育関係者には、こうした分析手法は必ずしも馴染みのある概念ではなかった。このため、「科学的根拠のある研究」をどのように解釈し運用するか、教育現場に戸惑いが生じることとなった。

こうした状況を受け、教育省は、2003年12月に「厳密なエビデンスによって裏付けられた教育実践の識別と実施：ユーザーフレンドリー・ガイド」と題したガイドラインを発表した(Coalition for Evidence-Based Policy 2003)。このガイドラインでは、教育にかかわる介入(intervention)の有効性を、どのように見極めるかを説明している。

それによれば、最も厳密なエビデンスが得られるのは、ランダム化比較試験(Randomized Controlled Trial: RCT)である。生徒やクラス、学校などをランダムに実験グループと比較グループに分けるこの方法によれば、他の要因による影響を排し、純粋な介入の効果を測定することができるとされた。教育分野でのRCTの適用例としては、学力面で遅れをとるリスクのある子供に対する個別指導、学級規模の縮小、英語習得のための発音練習などを対象としたものがあり、それぞれ有効性が確認されている。

教育分野でより広く使われているのは、ランダム化以外の方法で比較グループを設定する擬似実験デザインである。この方法では、グループ間に存在する見かけではわからない差異のために、正しい結果が得られない場合がある。同ガイドラインは、RCTと擬似実験デザインで正反対の結果が

出たケースを紹介し、この方法だけで最終的な結論を出すことはできないとしている。さらに、学校現場でしばしば行われる事前事後比較では、意味のある評価結果は得られないとされた。比較グループがないままに施策を適用し、実施前と実施後のデータを比較しても、その変化が施策によるものなのか、施策が無くても生じたものなのかを区別できないためである。

このように、教育省は、施策のエビデンスとしてRCTを最も上位に位置づけている。RCTは、医学や農学などの自然科学で一般的に用いられてきた手法である。それを教育分野にも採用し、施策の有効性をあらかじめ実験的に確かめることが意図されている。

(3) RCTをめぐる論争

このRCTの重視は、教育や評価にかかわる研究者の間に、大きな論争を巻き起こすこととなった。その論点は多岐にわたるが、ここでは二つの主要な論点に絞って紹介する。

第一の論点は、RCTを他の評価手法より上位に位置づけることの是非である。一部の評価研究者からは、擬似実験や相関分析などでも因果関係の検証は可能であり、RCTが現実社会の複雑な因果関係を検証するのに適さない場合もあるなどの反論が提起された (Donaldson and Christie 2005, p.63)。こうした見方は、上記の教育省の立場と真っ向から対立する。

第二の論点は、教育現場でRCTの広範な適用が可能かどうかという問題である。社会科学におけるRCTの実施は、倫理的に難しい場合があることが指摘されている。加えて、学校や生徒を取り巻く状況は多様かつ複雑であるため、医療分野で行われるような統制された実験は難しく、実験結果の一般化もしにくいという議論がある (Glenn 2004)。これに対し、RCTを擁護する立場からは、医療においても患者の個人差や症状の多様性を織り込んだ上で一般化可能な実験を行っているという反論が示されている (Slavin 2004)。

多くの場合、RCTに批判的な論者であっても、必ずしもRCTの有用性自体を否定しているわけではない。RCTは強力な評価手法ではあるが、現

実に適用できるケースは限られており、RCT以外の研究にも重要な役割があるというのがその主張である。反対に、RCTを擁護する論者も、RCT以外の手法の有用性を否定しているわけではない。NCLB法のいう「科学的根拠のある研究」には様々な研究が含まれるが、施策の有効性、すなわち "What works?" という問いに答えるためには、RCTが最も有用だというのがその主張なのである (Mageau 2004, p.2)。

アメリカの評価関係者の間には、実験デザインに代表される定量的評価の優位性を主張する立場と、定性的評価を重視する立場とがあり、長い間、論争を繰り返してきた歴史がある (龍・佐々木 2000, pp.12-14)。上記の意見対立には、こうした従来からのパラダイム論争が色濃く反映されている。いいかえれば、NCLB法を契機に初めてRCTをめぐる論争が巻き起こったわけではなく、専門家間で続いてきた大きな論争が、NCLB法によってあらためて表面化したということになる。

(4) 労働・福祉分野の経験

以上のような議論があるにもかかわらず、RCTの有用性が、ここにきて強調されているのはなぜなのだろうか。一つの重要な背景となっているのは、コクラン共同計画とEBM (Evidence-Based Medicine) の発展、キャンベル共同計画やイギリスにおけるエビデンス重視の動き (Evans and Benefield 2001) であろう。こうした一連の活動の中では、RCTは最もレベルの高いエビデンスを産出するものと位置づけられている。これに加えアメリカの場合、労働や福祉など他の社会科学分野において、RCTがすでに広く実施され、政策決定に影響を与えてきた経緯があることを指摘しておきたい。

アメリカで、社会科学におけるRCTの適用が始まったのは1960年代である。その時点では、現実社会における実験が果たして可能なのかどうか、また有用な情報が得られるのかどうかは未知数であった。それ以降、労働、社会福祉、刑事司法などの分野で、数多くのRCTが実施されてきたが、特に事例が多いのは、福祉政策が雇用に与える影

響や、職業訓練を対象としたRCTである (Greenberg and Shroder 2004)。この結果、RCTは十分に適用可能であり、政策決定に有用な情報を与えることがわかってきた。

RCTの結果が、政策決定に直接的な影響を与えた例の一つに福祉改革がある。1996年までの社会保障法では、州が新しい福祉改革アプローチを試行する際に評価を義務付けていたため、州の施策を対象とした数万人規模のRCTが実施された (Gueron 2002, p.33,42)。1988年には、福祉受給者の就職を支援する法改正が行われたが、そこでは、RCTによる評価結果が重要な役割を果たしたとされる (Szanton 1991)。

RCTは、評価結果の信頼性が高いことから、他の手法に比べ、政策決定への影響力が大きいと指摘されている (Burtless 2002, p.183)。たとえば、1980年代の労働省の職業訓練プログラムでは、擬似実験による評価結果にバラつきがあり、正しい施策の効果が判然としなかった。このため、専門家の助言を受けた労働省がRCTによる評価を行ったところ、若年層を対象とした職業訓練には効果がないことが判明し、この結果、効果が認められた他の年齢層に大きく予算が振り替えられた。RCTによる評価結果に説得力があったため、議会も抵抗なく行政府の決定に従ったという。

こうした経験を通じ、労働や福祉の分野では、RCTはもはや特別なものではなくっており、プログラム評価の一環として、ごく普通に実施されている。Abt Associates、Mathematica Policy Research、MDRC (Manpower Demonstration Research Corporationから2003年に改称) など、RCT実施のノウハウを持つコンサルタントも成長した。MDRCは、2002年までの時点で30の大規模なRCTを実施しており、200以上の地域において30万人近い人々を対象としたという (Gueron 2002, p.16)。

これに対し教育分野では、RCTが行われる場合も、施策の有効性評価よりは理論的な検証が多く、実験期間も短かったと指摘されている (Slavin 2002, p.16)。教育施策のRCTの事例としては、1960年代に早期教育の有効性を評価したペリー就学前教育プロジェクトや、1980年代に学級規模と学力の因果関係を検証したテネシー州STARプロ

ジェクトなどが有名である。しかし、全体として見れば実施は限定的であった。

労働・福祉分野と教育分野で、このような違いが生まれた背景として、次のような要因が指摘されている (Burtless 2002 ; Gueron 2002)。まず、労働省や保健福祉省では、経済学者など厳密な評価に積極的な専門家が影響力ある地位に就いたが、教育省では、そうした状況が生まれなかった。また、福祉では連邦政府の政策への影響力が大きいのに対し、教育は原則として州の所管であり、連邦の影響力が小さかったことがある。連邦のほうの評価予算や人材面で恵まれているため、連邦の影響力が大きい分野のほうが、RCTが実施されやすい傾向がある。さらに、労働や福祉分野では、短期間で実施できる介入が多いこと、実験グループによる比較グループへの波及的な影響を遮断しやすいことなど、教育と比較して実験の実施に適していた要素もあった。

果たして、このような過去の傾向をくつがえし、教育分野でのRCTの産出と利用は進むのだろうか。その答えは、今後、実際にNCLB法を運用していく中で検証していくほかはないように思われる。そして、NCLB法の成否の一つのカギを握っているのが、次に述べるWWCの動向である。

2. What Works Clearinghouse (WWC) の概要と現状

(1) 目的と位置付け

NCLB法によって、教育界には、エビデンスを求める強い需要が生まれた。しかし、膨大な教育分野の研究の中から、質の高いエビデンスを見分けることは容易ではない。NCLB法を運用していくためには、科学研究をわかりやすい言葉で教育関係者に伝える仕組みが必要となる。この役割を期待されているのが、WWCによる系統的レビューである。

WWCは、教育にかかわる政策、施策、活動、製品の有効性について、最新で質の高いレビューを提供し、エビデンスに関する信頼できる情報源となることを目的としている。2004年7月には、

第一弾のレビュー結果を掲載したウェブサイトが公開された。WWCの運営は、教育省の教育科学研究所が、American Institute for Research とキャンベル共同計画によるジョイント・ベンチャーへの契約を通じて行っている。

(2) 系統的レビューのプロセス

WWCにおける系統的レビューのプロセスは、以下のとおりである。

レビュー・トピック

まず第一に、レビュー対象となるトピックが選定される。WWCは、ウェブサイトや専門家のネットワークを通じ、トピックの候補を募集している。選定にあたっては、生徒への成果が期待できるか、現場での適用可能性があるか、教育関係者にレビューに対する需要があるか、科学的研究が存在しそうかどうか、などが基準とされる。表1に、WWCがこれまでに選定したレビュー・トピックをまとめた。

表1 WWCで選定されたレビュー・トピック

- ・ 中学校の数学カリキュラム
- ・ 小学3年生までの初期の読解
- ・ 人格（道徳）教育
- ・ 高校の退学防止策
- ・ 英語を母語としない小学生の英語学習
- ・ 小学校の算数カリキュラム
- ・ 3～5歳児の就学前教育
- ・ 中学校・高校の非行・暴力的行動の防止策
- ・ 成人の読解力向上策
- ・ 小学生同士の相互指導とグループ学習

(出所) WWC ウェブサイトをもとに
筆者作成 <http://w-w-c.org/topics/current_topics.html>

レビュー・チーム

続いて、あるトピックを担当するレビュー・チームが編成される。それぞれのチームには、当該分野の専門家と分析手法の専門家がアドバイザーとして含まれる。彼らの指導のもとで複数のリサーチ・アナリストがレビューを行い、プロジェクト・コーディネーターがレビュー・プロセス全体の監督を行う。

文献検索とスクリーニング

レビュー・チームはまず、体系的かつ包括的な文献検索を行い、当該トピックにかかわる研究を収集する。検索対象には、査読誌だけでなく、非査読誌や非出版物も含まれる。教育関係者や出版社などが、関連する研究を推薦することもできる。集められた研究は、トピックとの関連性や新しさなどの観点からスクリーニングされる。

エビデンス基準

続いてレビュー・チームは、個々の研究のエビデンスとしての信頼性をWWCエビデンス基準に基づいて判定する。その結果、それぞれの研究は、以下の3カテゴリーにランク付けされる。

「エビデンス基準を満たす」...RCT、回帰非連続デザイン

「留保つきでエビデンス基準を満たす」...擬似実験デザイン、ランダム化・サンプルの脱落・実施過程等に問題のあるRCT、サンプルの脱落・実施過程等に問題のある回帰非連続デザイン

×「エビデンス・スクリーンを通らない」...上記以外の研究

「エビデンス基準を満たす」、「留保つきでエビデンス基準を満たす」と判定された研究は、介入の正確さ、アウトカム指標、統計分析などの重要な側面について、さらなるレビューを受ける。

エビデンス・レポート

レビューの結果は、WWCの最終成果物であるエビデンス・レポートにまとめられる。これに含まれるのは、「エビデンス基準を満たす」、「留保つきでエビデンス基準を満たす」と判定された研究のみである。エビデンス・レポートには、個々の研究の要約である「研究レポート」、介入ごとのエビデンスをまとめた「介入レポート」、トピック全体のレビュー結果を総合した「トピック・レポート」の3種類がある（表2）。WWCは、新しい研究結果を含めるために、エビデンス・レポートを定期的にアップデートする予定である。

表2 WWCのエビデンス・レポート

レポート	内 容
研究レポート	エビデンス基準を満たす個々の研究を要約したもの。当該研究の概要、研究結果、エビデンス基準を満たすかどうか、研究の長所と短所などが記載される。
介入レポート	エビデンス基準を満たす研究（留保つきであってもよい）が一つ以上ある介入について作成される。当該介入についての説明、個々の研究結果の要約、全体としてのエビデンスの強さなどが記載される。付録として、個々の研究から得られた統計データの概略が掲載される。
トピック・レポート	当該トピックに関連するすべてのレビュー結果をまとめたもの。エビデンス基準を満たす全研究結果が要約され、総合的なエビデンスの状況が説明される。

(出所) WWC ウェブサイトをもとに掲載されているエビデンス・レポートをもとに筆者作成
http://w-w-c.org/topics/current_topics.html

品質管理

WWCには、リサーチ・デザインや評価研究の専門家からなるテクニカル・アドバイザー・グループ（TAG）が儲けられている。レビュー・プロセスのあらゆる段階で、このTAGやその他の専門家によるピア・レビューが行われる。また、レビューを受けている研究の実施者や、レビュー対象の教育ソフトウェアの開発者などにも、エビデンス・レポートをレビューする機会が与えられる。最後に、教育省がレビュー・プロセスが適切だったかを検証し、承認されたものがウェブサイトにアップされる。

(3) 系統的レビューの実際

現時点で、WWCのウェブサイトに掲載されているエビデンス・レポートは、「中学校の数学カリキュラム」というトピックに関するもののみである。ここでは、その概要を紹介する。

アメリカでは、国際的な学力調査でアメリカの子どもの数学力が低位に甘んじていることもあり、数学力を向上させねばという問題意識が強い。そのための方策の一つとして、画一的な内容だった従来の教科書に代わる、新しいタイプの数学カリキュラムの開発が進んでいる。新種の教材、コンピュータ・ソフト、ビデオなどを用い、指導方法を工夫することによって、生徒の数学的理解を深めることを意図したものである。こうした数学カリキュラムは、すでに一部で普及が始まっているが、その有効性については専門家の間でも見解が分かれていた。

WWCは、様々な情報源をもとに、中学生を対象とした44種類の数学カリキュラムを特定した。このうち、関連する評価研究が存在した20種類のカリキュラムに対するレビュー結果が、WWCウェブサイトに掲載されている（表3）。これらカリキュラムに対する評価研究のうち、文献検索の段階でのスクリーニングを通過したものは78件であった。それらがエビデンス基準に基づいてレビューされた結果、4件が「エビデンス基準を満たす」、6件が「留保つきでエビデンス基準を満たす」と判定され、残る68件は「エビデンス・スクリーンを通らない」とされた。この結果、10件の研究レポートと5件の介入レポートが作成された。表4に、介入レポートの要約表の一例を示す。

トピック・レポートは、以上のレビュー結果を総合したものである。エビデンス基準を満たす研究が存在したのは、表5に示す5種類のカリキュラムである。このうち、統計的に有意な有効性が示されたのは、Cognitive Tutor®とI CAN Learn® Mathematics Curriculumの2種類であった。それぞれについて有効性を示す研究は1件のみである。こうした結果から、トピック・レポートでは、中学校の数学カリキュラムについてのエビデンスは貧弱であるとして、評価手法の改善を呼びかけている。

ここで注意してほしいのは、ある介入について十分なエビデンスがあるということと、その介入が有効であるということとは別だという点である。信頼性の高い研究が存在したとしても、その結果が「当該介入は有効でない」という場合もある。同様に、ある介入についてのエビデンスが存

表3 トピック「中学校の数学カリキュラム」のレビュー結果

カリキュラムの名称	エビデンス基準			介入レポートの有無
	満たす	留保つきで満たす	満たさない	
Adventures of Jasper Woodbury Series	0	0	1	-
Algebra Project	0	0	2	-
Algebraic Thinking	0	0	1	-
Cognitive Tutor®	1	0	5	
CompassLearning	0	0	2	-
Connected Mathematics Project (CMP)	0	3	12	
Connecting Math Concepts (CMC) mathematics program	0	0	2	-
CORD Applied Math	0	0	1	-
Destination Math	0	0	1	-
I CAN Learn® Education Systems	1	2	4	
Integrated Mathematics, Science, and Technology (IMaST)	0	0	1	-
Mathematics in Context (MiC)	0	0	3	-
MATHThematics	0	0	3	-
Model Mathematics Program (MMP)	0	0	1	-
Moving with Math	0	0	12	-
Partnership for Access to Higher Mathematics (PATH Mathematics)	0	0	1	-
PLATO	0	0	1	-
Saxon Math	1	1	4	
Successmaker	0	0	9	-
The Expert Mathematician	1	0	0	
Studies not categorized by intervention	0	0	2	-
合計	4	6	68	
レポート数	4	6	-	5

(出所) WWC ウェブサイトをもとに筆者作成 <<http://w-w-c.org/topics/TopicInterventions.asp?EvidenceRptID=03>>

在しないということは、その介入が有効でないことを意味しない。まだ科学的に有効性が検証されていないだけで、今後の研究しだいで有効性が明らかになる可能性もある。

(4) WWCと意思決定

冒頭でも述べたように、WWCの場合、エビデンスの使い道がNCLB法によってあらかじめ用意されている点が、キャンベル共同計画やコクラン共同計画との大きな違いである。今後、WWCで収集され、レビューされたエビデンスが、どのように教育現場で活用されていくかが注目される。

教育省の関係者は、WWCはあくまで意思決定を助ける情報を提供するものであり、特定の施策や製品を支持するものではないという点を強調し

ている (Megeru 2004, p.35)。NCLB法はエビデンスに基づいた施策の採用を求めているが、実際にどのような施策を科学的根拠があると判断するかは現場に委ねられており、WWCのエビデンス基準を満たす施策のみを採用せよということではない。つまり、WWCはあくまで参考情報であるというのが教育省の立場である。

しかし、教育現場は、WWCをもっと拘束力のあるものと捉え、過剰に反応するおそれも指摘されている。これは、連邦の補助金を受ける際に、特に問題となる。NCLB法のいくつかの条文では、連邦の交付する資金を「科学的根拠のある研究」で有効性が確認されたものだけに使用するように求めている。このため、現場が補助金を受けられなくなることを恐れ、WWCの基準を満たさない施策の採用を自粛する可能性が出てくる。WWCが

表4 I CAN Learn® Mathematics Curriculumの介入レポート要約表

介入	I CAN Learn® Mathematics Curriculumは、基礎代数・代数のための双方向のソフトウェア・システムである。初歩的な数学から高等代数までの326のレッスンから成る。生徒が自分のペースで学習するためにデザインされており、主に7年生から10年生の人種的に多様なインナー・シティの生徒を対象としている。
対象	中学生
評価結果	<ul style="list-style-type: none"> ・1件のRCT（1校、254人）によれば、I CAN Learn を利用した生徒は、州の数学テストにおいて、通常のカリキュラムに従った比較グループよりも統計的に高い点を取ったことが示された。 ・1件の擬似実験によれば、I CAN Learn を用いた生徒が比較グループより高い点数をとったが、その差は統計的に有意ではなかった。 ・もう1件の擬似実験によれば、I CAN Learn を使った生徒が比較グループより高い点数をとったが、分析に欠陥があるため、結果の有意性を正確に見極めることができない。
エビデンスの基盤	エビデンス基準を満たす...1（RCT） 留保つきでエビデンス基準を満たす...2（擬似実験） × エビデンス・スクリーンを通過しない...4
エビデンスの限界	I CAN Learn® Mathematics Curriculumのエビデンスの基盤は3件の研究であり、そのうち1件のみがRCTによるものである。その他は擬似実験によるもので、2グループ間の測定されていない差異が結果に影響を与えている可能性があり、エビデンスとしてはRCTより弱い。擬似実験のうちの1件は、推計されたインパクトの信頼区間を計算していないという分析上の欠陥がある。
利用状況	このカリキュラムは1995年に最初に導入された。2004年10月までに、23州501校で利用されている。典型的な利用例は、タイトル ^(注) 指定校や、都市部の大規模学区における学力面で遅れをとるリスクのある生徒、マイノリティーの生徒である。
開発者と連絡先	JRL Enterprises, Inc., 400 Poydras Street, Suite 100, New Orleans, LA 70130; telephone (504) 263-1380; www.icanlearn.com.

（注）タイトル とは、社会経済的に恵まれず学力が低位の生徒に対し、連邦から補助金を支給する施策である。

（出所）Institute of Education Sciences, *What Works Clearinghouse Intervention Report: I CAN Learn® Mathematics Curriculum*. Updated December 1, 2004, p.1

どの程度、現場へのインパクトを持つかは、地域、施策内容、分野により異なると予想されている（Schneiderman, 2004, p.32）。

（5）エビデンスの需要と供給

現在のところ、教育分野でのRCTの事例は非常に限られており、エビデンスに対する需要が、供給を大きく上回っている状況にある。中学校の数学カリキュラムのレビューでも、WWCが特定した44のカリキュラムのうち、24については、評価が一件も行われていなかった。この需要と供給の大きなギャップは、エビデンスを積極的に生み出させる強いインセンティブとなる。

教育省は、プログラム評価を実施する際に、実験デザインによる研究に優先順位を与える方針を発表している⁴。この結果、2004年に行われた同

省の施策の有効性評価のうち、90%でRCTが採用されている（U.S. Department of Education 2005）。その中には非常に大規模なものもあり、たとえば、Mathematica Policy Researchは、16種類の商業ソフトウェアの英語と数学への効果を評価するために、560万ドルの資金を受けている（Glenn 2004）。

評価研究への需要と同時に、評価研究のレビューに対する需要もまた増加している。もともとRCTの実施例が限られていることを反映し、WWCがレビュー対象に選定しているトピックはまだ限定的なものである。WWCによるレビューがさらに進展するまでの間は、教師や地方政府の行政官など、教育関係者自身に、エビデンスを見極める目が求められる。WWCのレビュー・トピックの中でも、英語読解などはRCTが比較的多く存在する分野であり、そのレビュー結果が待たれるところである。

表5 トピック「中学校の数学カリキュラム」に関するエビデンスの要約

介入	サンプルと環境	エビデンスの強さ		
		研究の強さ	サンプル数	効果の推計結果
Cognitive Tutor®	・ 9年生	RCT	<u>1RCT</u> ・ 360人	<u>RCT</u> ・ Percentile rank: 9.1 ・ Standard scores: 0.23
Connected mathematics Project	・ 6-8年生 ・ 中西部、西部、東部、北東部、南部 ・ 農村部、都市部、郊外 ・ 社会経済状況が高位の地域と低位の地域	擬似実験	<u>3擬似実験</u> ・ 1,095人 ・ 50校 ・ 36-42校	<u>擬似実験</u> ・ 有意ではない
The Expert Mathematics Curriculum	・ 8年生 ・ 低所得層	RCT	<u>1RCT</u> ・ 70人	<u>RCT</u> ・ 有意ではない
I CAN Learn® Mathematics Curriculum	・ 7-10年生 ・ 都市部のタイトル指定校	RCT 擬似実験	<u>1RCT</u> ・ 254人 <u>2擬似実験</u> ・ 169学級 ・ 116学級	<u>RCT</u> ・ Percentile rank: 15.9 ・ Standard scores : 0.41 <u>擬似実験</u> ・ 有意ではない
Saxon Math	・ 8年生 ・ 都市部	RCT 擬似実験	<u>1RCT</u> ・ 36人 <u>1擬似実験</u> ・ 78人	<u>RCT</u> ・ 有意ではない <u>擬似実験</u> ・ 有意ではない

(出所) Institute of Education Sciences, *What Works Clearinghouse Topic Report: Curriculum-based interventions for increasing K-12 math achievement - middle school*. Updated December 1, 2004, p.4

今後、アメリカの教育界で、どれだけRCTが生み出され、レビューされ、活用されていくのか。WWCの息の長い取り組みは、ようやく動き始めたばかりである。

おわりに～日本への示唆～

今、わが国では、学校評価への関心が高まっている。それぞれの学校が、目標を設定した上で教育活動を評価し、その結果を教育の改善に役立てることをねらったものである。しかし、学校を評価するという発想はあっても、学校で行われる施策や指導方法の一つ一つを評価するという発想は、希薄なように感じられる。日本の教育現場でも、習熟度別授業、ITの活用、学級規模の縮小など、多くの熱意ある取り組みが行われている。こうした取り組みが、学力向上など意図した効果を上げ

ているか科学的に評価し、効果のある取り組みを採用していこうというのがNCLB法の発想なのである。

NCLB法は、まだ緒に付いたばかりであり、その成否は今後の検証を待たなければならない。また、RCTを特に重視することに関しては、アメリカにおいても議論のあるところである。しかし、エビデンスに基づく教育を目指すNCLB法の方向性は、わが国にとっても示唆に富むものといえるだろう。

数年前に話題になった学力論争が浮き彫りにしたのは、日本では、そもそも議論のベースとなる学力のデータが十分存在せず、学習指導要領の変更などの重要な政策変更が、科学的に評価されていないという問題であった。定性的な評価と定量的な評価はいずれも大切なものであるが、わが国の場合、定量的な評価が不足している感は否めない。今後はわが国でも、定量的な有効性評価の蓄

積が必要といえよう。アメリカのようなRCTの実施が日本でどの程度可能かどうかは議論の余地があるが、大切なのは、それぞれの施策に応じ、適切なデザインを用いた評価を行うことである。その上で、評価の質をレビューし、意思決定につなげるためのインフラとして、WWCのような系統的レビューが求められる。文部科学省は、学力向上フロンティア事業で各県の実践研究を支援しているが、こうしたモデル事業の評価から始めるのも一案であろう。

注記

- 1 <http://w-w-c.org>
- 2 正式名称An act to close the achievement gap with accountability, flexibility, and choice, so that no child is left behind. P.L. 107-110.
- 3 実験的手法には、ランダム化バイアス、代替バイアスなどの手法上の問題があることが指摘されているが（労働政策研究・研修機構 2004、p.11）アメリカの議論においては、こうした点は主たる論点とはなっていない。
- 4 Department of Education, RIN 1890-ZA00, " Notice of proposed priority," *Federal Register*, 68 (213), Nov. 4, 2003 ; " Notice of final priority," *Federal Register*, 70(15), Jan. 25, 2005. 教育分野では、従来、定性的な研究が主流であったため、評価予算の一部は、労働経済学者などRCTを頻繁に行ってきた他分野の研究者に流れているという（Glenn 2004）。この評価予算の配分問題は、RCTをめぐる研究者間の論争にも影響を与えている。

参考文献

- 教育政策の評価手法研究会（2004）『プログラム評価の手法とその活用』（平成15年度文部科学省委嘱研究：教育行政の評価手法の在り方に関する調査研究報告書）
- 松山幸弘（2001）「アメリカの教育改革から学ぶこと」、『*Economic Review*』、5（4）：93-109
- 龍慶昭・佐々木亮（2000）『「政策評価」の理論と技法』、多賀出版
- 労働政策研究・研修機構（2004）『労働市場政策の効

果に係る定量的評価の欧米における先行研究についての調査』（労働政策研究報告書 No.L-4）

- Boruch, R., de Moya, D., and Snider, B. (2002). The importance of randomized field trials in education and related areas. In F. Mosteller & R. Boruch (Eds.), *Evidence Matters: Randomized Trials in Education Research*. Washington, DC; Brookings Institution Press.
- Burtless, G. (2002). Randomized field trials for policy evaluation: why not in education? In F. Mosteller & R. Boruch (Eds.), *Evidence Matters: Randomized Trials in Education Research*. Washington, DC; Brookings Institution Press.
- Carlson, J. S., and Levin, J. R. (Eds.). (2005). *The No Child Left Behind Legislation: Educational Research and Federal Funding*. Greenwich, Conn: Information Age Publishing.
- Coalition for Evidence-Based Policy. (2003). *Identifying and Implementing Educational Practices Supported by Rigorous Evidence: A User Friendly Guide*. Washington, DC: U.S. Department of Education, Institute of Education Sciences.
- Cook, T. D., and Payne, M. R. (2002). Objecting to the objections to using random assignment in educational research. In F. Mosteller & R. Boruch (Eds.), *Evidence Matters: Randomized Trials in Education Research*. Washington, DC; Brookings Institution Press.
- Donaldson, S. I., and Christie C. A. (2005). The 2004 Claremont Debate: Lipsey vs. Scriven. *Journal of MultiDisciplinary Evaluation*, 60-77.
- Evans, J., and Benefield, P. (2001). Systematic reviews of educational research: does the medical model fit? *British Educational Research Journal*, 27 (5), 527-541.
- Fletcher G. H. (2004). Scientifically based research: guidelines or mandates for product purchasing? *T.H.E. Journal*, 31 (8), 22-25.
- Glenn, D. (2004). No classroom left unstudied. *The Chronicle of Higher Education*, 50 (38), A12-A15.
- Greenberg, D., and Shroder, M. (2004). *The Digest of Social Experiments, 3rd Edition*. Washington, DC: Urban Institute Press.
- Gueron, J. M. (2002). The politics of random assignment: implementing studies and affecting policy. In F. Mosteller & R. Boruch (Eds.), *Evidence Matters: Randomized*

- Trials in Education Research*. Washington, DC; Brookings Institution Press.
- Krueger, A. B. (1999). Experimental estimates of education production functions. *The Quarterly Journal of Economics*, 114, 497-532.
- Mageau, T. (2004). Determining 'what works'; an interview with Dr. Grover 'Russ' Whitehurst. *T.H.E. Journal*, 31 (6), 32-37.
- Mervis, J. (2004). Meager evaluations make it hard to find out what works. *Sciences*, 304, 1583.
- Schneiderman, M. (2004). What does SBR mean for education technology? *T.H.E. Journal*, 31 (11), 30-36.
- Slavin, R. E. (2002). Evidence-based educational policies: transforming educational practice and research. *Educational Researcher*, 31 (7), 15-21.
- Slavin, R. E. (2004). Education research can and must address "what works" questions. *Educational Researcher*, 33 (1), 27-28.
- Szanton, P. L. (1991). The remarkable "quango": knowledge, politics, and welfare reform. *Journal of Policy Analysis and Management*, 10 (4), 590-602.
- U.S. Department of Education (2002). *No Child Left Behind: A Desktop Reference*.
- U.S. Department of Education (2005). *FY2006 Program Performance Plan*.
- Viadero, D. (2004). Ed. Dept. issues practical guide to research-based practice. *Education Week*, 23 (16), 12. (2006.2.16受理)

Evidence-Based Education What Works Clearinghouse in the United States

Satoko Tanabe

National Diet Library
s-tanabe@ndl.go.jp

Abstract

The No Child Left Behind Act, which was enacted in the United States in 2002, includes challenging provisions that require states and local schools to adopt programs and practices based on scientifically-based researches. The act focuses especially on studies using experimental designs. In order to provide the education community reliable information on "what works," Department of Education opened a web site called What Works Clearinghouse (WWC). WWC review teams conduct systematic reviews of studies and the results are disseminated as WWC evidence reports. WWC is now under development, and is likely to influence U.S. education programs by its contents.

The idea of evidence-based education is suggestive for the Japanese education community. It is desirable that we accumulate reliable evaluation studies and conduct systematic reviews.

Keywords

evidence, program evaluation, experimental designs, education, What Works Clearinghouse

【実践・調査報告：依頼原稿】

ODA分野における『エビデンスに基づく評価』の試み： 「貧困アクションラボ」の動向

佐々木 亮

ウェスタンミシガン大学

ryoh.sasaki@wmich.edu

要 約

医療分野のコクラン共同計画、社会政策分野のキャンベル共同計画に大きな影響を受けて、ODA分野でも厳格な実験デザインを適用した一次評価を産出し、将来的にはデータベース化して提供することを目指す「貧困アクションラボ」というプロジェクトが始まっている。その動向を踏まえつつ、実験デザインを巡る諸問題である倫理の問題、費用の問題、バイアスの問題、準実験デザインではなく実験デザインが望ましい理由、そして、今まで実験デザインがODA分野でほとんど使われなかった理由を議論する。さらに、日本の当該セクターの現状と課題を議論し、最後に、ODA分野でも実験デザインによるインパクト評価を行っていくこと、及び同プロジェクトを資金的に支援している世界銀行の最大の出資者である日本が同プロジェクトに注目し、より積極的に関与していくことを提言する。

キーワード

実験デザイン、ODA、要請主義、貧困アクションラボ、Poverty Action Lab

1. 組織的活動の起源、背景、問題意識

ODA分野は、世界的に見ても比較的早くから評価が実施されてきた公共分野である。例えば、評価のもっとも初期的な文献とされるハイエスの評価に関するモノグラフは「開発プロジェクトの評価」という題であった（Hayes, 1959）。また、評価に関するテキストとしてもっとも広範に利用されている「評価：体系的アプローチ」（Rossi, Freeman, Wright, 1979）は、もともとユネスコの会議で発表された二つの別々の論文がもとになっ

ている（Rossi et al, 1993, p.x）。ロッシとライトが書いた論文と、フリーマンが書いた二つの論文は、それぞれユネスコの依頼に基づいて、途上国の開発プロジェクトの評価に利用できる手法をレビューした論文であった。3人はそれをもとに加筆して、開発プロジェクトの評価手法を一冊の本にまとめたが、それが「評価を行う」（Rossi, Freeman, Wright, 1980）であった。そして同書をもとに、アメリカ国内の公共政策一般を対象にして作成されたのが前出の「評価：体系的アプローチ」（Rossi et al, 1979）であった。したがって、もと

もとODA分野のために編纂された本がアメリカ国内向けに転用されたと見ることができるのだ。

ところで、「評価：体系的アプローチ」の初版 (Rossi et al, 1979) は、キャンベルの「リサーチのための実験と準実験デザイン」(Campbell, Stanley, 1966) の影響を強く受けて、実験デザインと準実験デザインの解説に過半のページを割いていた。この伝統は版を重ねて受け継がれ、途中の版では各種の準実験デザインを追加して構成が煩雑になったが、同書の最新版である第7版では、再び、実験デザインと準実験デザインというシンプルな章立てとなった (Rossi et al, 1982, 1985, 1989, 1993, 1999, 2004)。

ODA分野において、理論としては広く知られていた実験デザインが実際に最初に用いられた事例が何であるかは定かではない。1970年代に、コロンビアの児童の知覚開発プロジェクト (1971-1975年、USAID支援) の評価と、ニカラグアの遠隔教育プロジェクト (1974-1978年、フォード財団支援) の評価で適用されたとの記述があり (Rawlings, 2003) これらが最も初期的な案件のひとつと言えるであろう。

ただし、1970年代から、途上国の開発プロジェクトの評価で実験デザイン (別名：ランダム化比較実験、実験計画法、ランダム実験モデル等) が広範に用いられてきたのかというと、保健分野を除けば、適用案件に関する入手可能な情報は限られる。世界銀行のオペレーション評価部 (Operations Evaluation Department) は、1970年代後半から評価を実施しているが、対象はODAプロジェクトの計画部分 (Project design) と実施部分 (Implementation) が主であった (Valadez, Bamberger, 1994, p.227)。その状況から1980年代後半に至って、より長期的なインパクト (介入によるネットの効果) の測定に重点が移っていったが、実際に利用が検討された主なインパクト評価手法は準実験デザインであった。さらに、それさえも「手法的に厳格すぎる」として非現実的だとみなされ (Valadez et al, 1994, p.228) 簡便で経済的なインパクトの測定方法が模索されて実際に手法が開発され普及していった (例えば USAID, 1987)。

こうした趨勢の中で、社会セクターにおいて近

年実施された実験デザイン適用の試みとしては、1990 - 1992年に実施されたフィリピンの中初等教育プロジェクトの評価がある (Tan, Lane, Lassibille, 1999)。また近年の別の試みとして、1993年から1997年に実施されたボリビアの社会投資基金プロジェクトの評価があげられる (Rawlings, 2003)。ただし、いずれも世銀によるパイロットプロジェクトとして小規模に実施されたに過ぎない。そして、簡便で経済的なインパクトの測定方法が主流をなす時代の趨勢の中で、これらは例外的な試みとして認識されていたと言える。

その後、2000年代に入って、こうした状況に大きな変化が訪れた。マサチューセッツ工科大学のジャミール (Abdul Latif Jameel)、ハーバード大学のバナージェ (Abhijit Banerjee) 等が中心となって設立した「貧困アクションラボ」(Poverty Action Lab) が、その変化の原動力である。このプロジェクトは、医療分野のkokラン共同計画、社会政策分野のキャンベル共同計画に大きな影響を受けて始められたものであり、20世紀に医療分野で実験デザインが革新的な役割を担ったのと同じように、ODA分野で実験デザインが革新的な役割を担うことを目指すとしている (Kremer, 2005, p.10)。

2. 貧困アクションラボの概要

同ラボは、2003年に設立された。医療分野におけるkokラン共同計画、社会政策分野におけるキャンベル共同計画と同様に、ODA分野において実験デザインを用いた評価結果をデータベース化して提供することを目指していると思われる。ただし、現在までのところ、実験デザインを利用した評価事例がODA分野にはほとんどないことから、同ラボに加盟している研究者が、同ラボに依頼された評価案件を実施して一次評価を産出している段階である。したがって、構造化抄訳や、そのためのプロトコル (システマティック・レビューの手続き) が整備されるまでには至っていない。

一方で、2003年設立でまだ2年しか経っていないにも関わらず、すでにのべ42件の評価に取り組んでおり、19件を終了して、23件が実施中である。

多くは世界銀行がファイナンスしているという特徴がある。また、アメリカの財団法人であるマッカーサー財団 (MacArthur Foundation) も積極的にファイナンスを支援しているのも特徴的である。対象国も、インド (8件)、インドネシア (2件)、フィリピン (2件)、ケニア (15件)、南アフリカ (2件)、マダガスカル (1件)、ペルー (2件)、コロンビア (2件) と、アジア、アフリカ、南米をカバーしている。(図1) その他、本来の対象ではなかったが、依頼に基づいてアメリカ国内で実験デザインを適用して実施した評価案件が、8件ある。分野も教育 (主に初等教育)、保健、ジェンダー、マイクロクレジット、地方分権化と多方面にわたっている。本論文の最後に42件の概要リストを添付したので参照されたい。(表3)

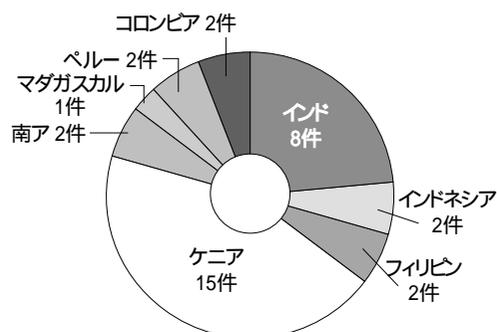
知識の共有 (データベース化) のための活動としては、ホームページを開設して、各評価案件の報告書を無料で公開している。また、「貧困と戦う：何が機能するか」(Fighting Poverty: What Works) と題する定期的な購読紙 (bulletin) を2005年中に発行開始する予定である。

今後、評価実施済みの案件が蓄積され、また同ラボ以外でも実験デザインを適用した評価結果が産出されるようになれば、コクラン共同計画、キャンベル共同計画のような体系的なデータベース化の作業が計画されると予想される。

3. 具体的評価事例

ODA分野は、その分野の中で、さらに教育、保健、地方行政、社会福祉などの小分野に分かれているわけであり、ODAの案件であるからと言って、当該案件の本質的な性格が先進国の国内の案件と変わるわけではない。ただし、先進国には一般に見られない途上国特有の状況があり、これに対しては、当該途上国のコンテキストの中で介入を検討し、その効果を検証する必要がある。次の案件は、ケニアにおける保健プロジェクトの例であり、世界銀行等の資金援助を受けて「貧困アクションラボ」が実験デザインを適用した評価結果の概要である。(ボックス1)

図1 評価案件の地域分布 (ODA分野)



(Source) Poverty Action Lab (2005) Projects by Status
<http://www.povertyactionlab.com/projects/>

4. 実験デザインに関する諸議論に対する回答の試み

同ラボの2003年の設立にあわせて、世銀のオペレーション評価部が主催してカンファレンスが開催された。同ラボの設立メンバーであるデュフロ (Esther Duflo) とクレマー (Michael Kremer) は、(1)なぜ実験デザインでなければならないか、(2)なぜデータベース化が必要か、(3)なぜ倫理的にも費用的にも問題ないと言えるのか、そして(4)なぜ今まで実験デザインがODA分野でほとんど使われなかったのか、に関して議論している (Duflo, Kremer, 2003)。以下では、まず彼らの分析を紹介し、さらに筆者の見解を加えて議論を深めた。

(1) なぜ実験デザインでなければならないか

回帰分析などの準実験デザインと、実験デザインを同一のODAプロジェクトに適用して、インパクトを評価したところ、評価結果が著しく違ったという研究結果がある (Glazerman, Levy and Meyers, 2002)。つまり、厳格な実験デザインによる評価では効果が確認されなかったにも関わらず、回帰分析では高い効果があると評価されたわけである。こうした状況はなぜ起こるのだろうか。クレマーによると、事後的な回帰分析では、インパクトとして、ネットの効果のほか、本来は含まれるべきではない測定エラーと、測定者のバイア

ボックス1 小学校における回虫駆除プロジェクト（ケニア）

世界の4人に一人は、腸内寄生虫に感染しており、それは途上国の小学生の間で特に顕著である。本プロジェクトはケニアのプシア県において、小学生に医療処置と関連教育を行うことを介入内容として、1998 - 2002年に実施された。同県の75校を、ランダム（無作為）に25校ずつ3つのグループに分けて、以下の年に介入を実施した（本来は3年連続の予定であったが、洪水が発生したので、2000年の分を2001年にずらして実施した）。

表1 各グループの介入実施年

	1998年	1999年	2000年	2001年
G1（25校）				×
G2（25校）	×			×
G3（25校）	×	×		

（注） × は介入実施を表す

このように時期をずらして実施することによって結局全ての学校が介入を適用されることになり、通常、実験デザインに関して指摘される倫理的な問題を回避している。

1998年末の時点では、G1を介入グループ、G2を比較グループとして比較できる（G3も比較グループとして利用できるが省略）。同時点のG1（介入グループ）の回虫感染率が27%で、G2（比較グループ）の同感染率が52%だったので、その差である - 25%が介入の効果であると評価された。

表2 介入結果

	1998年末
G1（25校）：介入グループ	27%
G2（25校）：比較グループ	52%
差（介入の効果）	- 25%

実験デザインを適用しているのも、もし同年に同県全域で何らかの感染症が大流行していたとしても、両グループ（G1とG2）が受ける影響は同程度になっているはずなので、上記の表に現れた - 25%という差は純粋な介入の効果であるとみなすことができる。

さらに、2001年末の時点では、介入を受けた学校の周辺の学校（G1～3以外の学校）でも、ある程度感染率が下がっていることが確認されたが、それは本介入による外部効果であると評価された。その外部効果を計算に入れると、合計で5,707感染数が削減されたことになる。

このほか、学校出席日数が、介入グループでは、比較グループと比較して15日間の増加（1-4学年）および10日間の増加（5学年以上）であった。従来は、保健指標だけで介入の効果を測定することが多かったが、学習指標にも効果が現れていることが確認された。なお、当該介入に要した費用と比較すると、一日あたりの出席を増加させるための費用はUS\$0.02と計算され、従来型の出席を増加させるための直接的なプロジェクトよりも、費用対効果が格段に優れていることが確認された。

（Source） Kremer, M., and Miguel, E. (2003) *Worms; Education and Health Externalities in Kenya*. Poverty Action Lab, MIT

スの2つがさらに含まれてしまうことになる。したがって簡単に言うと、本来は効果がないにも関わらず評価者が有するバイアスによって押し上げられた評価結果が出されることがしばしばあるとしている（Kremer, 2005, p.10）。この状況をアーバンインスティテュートのオール（Larry Orr）は、「フィッシング」（釣り）だと表現している。つまり、「標準的な回帰分析のテクニックを使って、自分が出したいと思う結果を出せるのであり、どの評価結果が一番正しいのかは誰も言うことができない」と述べている（Orr, 1999, p.xi）。こうした回帰分析の制約を克服するには、より厳格な準実験デザイン（回帰・分断デザインなど）が使用されるべきであるが、もっとも厳格なデザインが実験デザインであると言える。したがって、もっとも高い客観性を有するインパクト評価をしようとするれば、実験デザインを使用すべきということになる。ODA分野にもこの状況は当てはまるであろう。

ただし、彼らは、全ての評価で実験デザインが用いられるべきと主張しているわけではない。ODA分野では、簡便な評価が主流をなす一方で、実験デザインを用いた評価が今まで皆無に等しく、この状況は著しくバランスを欠いていると主張しているのだ（Duflo, Kremer, 2003, p.30）。

（2）なぜデータベース化が必要か

次に、データベース化の必要性であるが、デュフロとクレマーは「出版バイアス」（Publication Bias）の存在を指摘している。つまり効果があったという評価結果しか出版されない傾向があるため、出版を実現するために正の効果を出そうというバイアスが評価者にかかるというもので、分野を問わず一般にこうした出版バイアスは非常に甚大だという研究結果がある（DeLong, Lang, 1992）。これを回避するためには、データベースとして、出版されたものも出版されなかったものも登録しておくことが望まれる。また、組織として、効果がなかった、あるいはマイナスの効果があったという評価結果を受け入れる文化を持つべきであり、実際に医療分野ではこうした慣行が定着しているとされる（Duflo, Kremer, 2003, p.24）。ODA

分野でも統一された評価のデータベースの構築が望まれるが、現時点では、貧困アクションラボがその役割を担う可能性が高い。

（3）なぜ倫理的にも費用的にも問題ないと言えるのか

まず、倫理的な問題である。無作為割当（ランダム・アサインメント）によってサンプルとして参加した人たちの間に不公平が生まれるという主張がよくなされるが、介入適用に当たらなかった人が何か不利益を被っているわけではなく、もともとサンプルに加わらなかった人たちという母集団のレベルに戻されるだけである（佐々木、2003, p.107）。デュフロとクレマーは、この見解よりも若干控えめであり、無作為割当に当たらなかった人たちにも、一定期間後に介入を適用することによって倫理的な問題は回避できるとしている。事例として掲載したケニアの保健プロジェクトがまさにこれを実行している。また、通常、プロジェクト実施予算は限られているのが普通であり、毎期の予算制約の中で介入適用の順番を決める方法として、無作為割当が最も公平で、逆に倫理的だと主張している（Duflo, Kremer, 2003, p.20）。

また、実験デザインは通常、多大な費用を必要とすると言われるが、他のサーベイ調査となんら変わることはなく、必要な費用はそれ以上ではないと主張している。近年は比較的大規模なサーベイが実施されるケースが見られるようになったが、それならば実験デザインのみが高コストだと批判される所ではないことになる。また、その地域でしか有効でないサーベイ調査よりも、効果の有無を厳格に明らかにした評価結果という「国際的公共財」を実験デザインは提供するのであるから、費用に比べて利用価値が高いと主張している。そして、誤った公共政策を大規模に実施してしまうよりも、実験デザインによってその前に政策の誤りを明らかにする方が、よほど低コストで済むと述べている（Duflo, Kremer, 2003, p.20, p.28）。

(4) なぜ今まで実験デザインがODA分野ではほとんど使われなかったのか

この検討は、日本と世銀ではかなり状況が違うので注意が必要である。日本の特殊事情に関する検討は次節で扱うが、日本を含むODA一般を検証すると、案件選択過程の意思決定過程が大きく影響するとデュフロとクレマーは指摘している。例えば、前述の回帰分析などの準実験デザインによって、効果（例えば就学率の改善）が10%と推定されたプロジェクトと、厳格な実験デザインによって同2.5%と推定されたプロジェクトがあったとする。その場合に、予算配分を行う意思決定者としては、2.5%の評価結果となったプロジェクトを採用すべき理由がないとしている。また、多くの場合、意思決定者は複数の評価結果の質の違いを判断できなかつたり、過去の経験から推計された効果を割り引いて低く見てしまう場合も多い、とデュフロとクレマーは指摘している（Duflo, Kremer, 2003, p.29）。したがって、案件を提案してサポートする担当局の担当者も、予算配分を行う意思決定者も、実験デザインを適用して厳格に効果を測定するインセンティブはないということになる。こうした状況は現実を的確に指摘していると思われるが、実験デザインに関する不十分な理解に起因すると言わざるを得ず、実験デザインの利点が広く理解されることを待つしかない。

5. 日本の当該セクターの現状と課題

日本のODA分野で実験デザインが適用された評価結果はまだない。ただし、これはODA分野に限ったことではなく、日本の社会セクター各分野で実験デザインが適用された案件は極めて限られているとされる。日本では、医療分野、農業分野、教育心理学、工業生産管理等でその事例が観察されるだけである。ところが近年、ODAの開発調査に、いわゆる「実証」部分が導入されたわけで、このスキームを利用した実験デザインの適用の可能性が出てきているが、今までのところ事例はない。厳格にインパクトを特定しようという

発想が最もはっきりしていた案件は、インドネシアの初等教育案件（正式名称：インドネシア地域教育開発調査（REDIP）の実証部分であったと考えられるが、マッチングモデルの適用であった（同ファイナルレポート（2001））。

実験デザインの適用に関係すると思われる日本の特殊事情としては、援助の「要請主義」があげられる。これは、相手国政府自身が、自国の援助ニーズとそれに対する処方箋を熟知しており、それに基づいて最も合理的な援助要請を出してくるという前提に立っている。この前提が真実であれば、実験デザインの出る幕はない。しかし、実際には相手国政府が必ずしも最も合理的な処方箋を知っているわけではないため、援助国による知的支援が行われ、その結果に基づいて「要請を出させる」という一件矛盾した表現が非常に一般的に使われているのが実情である。また例え、要請主義の前提があるとしても、援助資金は日本の納税者の税金が投入されているわけであるから、相手国政府のみの判断ではなく、援助の効果が確かであると主体的に確認してから案件を実施することは、税金を投入する機関として当然になさねばならないことであろう。その責任を果たすために、実験デザインが適用できる余地があると思われる。

また日本の組織運営の原則にも改善の余地があると思われる。筆者の経験上、実験デザインを提案した際によく発せられるコメントは、「実験デザインを適用して、本当に効果がないとわかったら困る」というものである。これは、他の援助国・機関ではあまり見られない日本に特有の発想であると考えられるが、解決策としては、すでにデュフロとクレマーが指摘しているように、効果がなかった、あるいはマイナスの効果があったという評価結果を、組織として率直に受け入れる慣行を定着させるべきであるということに尽きる（Duflo, Kremer, 2003, p.24）。また、日本の組織においてはさらに、それを受け入れることが担当者に関するマイナスの人事評価にはならないという人事評価基準の変更が検討されるべきではないだろうか。

6. 日本の当該セクターへの教訓

日本は世界銀行の最大の出資者である。したがって、本論文で議論した世界銀行の資金的協力によって実施される「貧困アクションラボ」の実験デザインによる評価結果のデータベース化の試みとその成果には注目していくことが必要である。

また日本国内では、ODAは本当に役に立っているのかという意見が以前から聞かれていたが最近はさらに頻繁に聞かれるようになっている。こうした意見に対応して、実験デザインによるインパクト評価を行い、効果を確認し、そしてその評価結果を広報していくことは納税者の観点からも望ましいことである。また実験デザインによるインパクト評価は、ODAに関する意思決定者が行う合理的な意思決定にも貢献するであろう。

そして、コクラン共同計画の日本支部、キャンベル共同計画の日本支部が設立されて活発な活動や議論が行われているように、ODA分野においても「貧困アクションラボ」の日本支部が設立され、ODA分野では逆に本部をリードするくらいに成長することが期待される。

参考文献

Atkin, M. (2005). *Evaluation Roots*. Thousand Oaks, CA: Sage Publication.

Campbell, D.T., and Stanley, J.C., (1966). *Experimental and Quasi-Experimental Designs for Research*. Skokie, IL: Rand McNally.

DeLong, J. B., and Lang, K. (1992). Are all Economic Hypotheses False? In *Journal of Political Economy*, 100 (6) (December), 1257-72.

Duflo, E., and Kremer, M. (2003). *Use of Randomization in the Evaluation of Development Effectiveness*. Cambridge, MA: Poverty Action Lab, MIT.
<http://www.povertyactionlab.com/research/rand.php>

Glazerman, S., Levy, D. and Meyers, D. (2002). Nonexperimental Replications of Social Experiments: A Systematic Review. *Mathematica Policy Research, Inc. Interim Report/Discussion Paper*.

Kremer, M., - (2005). *Randomized Evaluations of Educational Programs in Developing Countries: Some Lessons*. Cambridge, MA: Poverty Action Lab, MIT.

Kremer, M., and Miguel, E. (2003). *Worms: Education and Health Externalities in Kenya*. Cambridge, MA: Poverty Action Lab, MIT.
<http://www.povertyactionlab.com/research/rand.php>

Hayes, S.P., Jr. (1959). *Evaluating Development Projects*. Paris: UNESCO.

Orr, L.L. (1999). *Social Experiments: Evaluating Public Programs With Experimental Methods*. Thousand Oaks, CA: Sage Publication.

IDCJ, PADECO (2001). *Regional Educational Development and Improvement Project (REDIP) Final Report*. Japan International Development Agency (JICA).

Rawlings, L.B. (2003). *Place Randomized Trials in Latin America*. Campbell Collaboration and Rockefeller Foundation.
http://www.campbellcollaboration.org/Bellaqio/Presentations/Rawlings_power.pdf

Rossi, P.H., Freeman.H.E., and Wright. J.D. (1979). *Evaluation: A Systematic Approach 1st edition*. Thousand Oaks, CA: Sage Publication.

Rossi, P.H., Freeman. H.E., and Wright, J.D. (1980). *Doing Evaluation*. Paris: UNESCO.

Rossi, P.H., and Freeman.H.E. (1982, 1985, 1989, 1993). *Evaluation: A Systematic Approach 2nd, 3rd, 4th, 5th edition*. Thousand Oaks, CA: Sage Publication.

Rossi, P.H., Freeman.H.E., and Lipsey, M. (1999). *Evaluation: A Systematic Approach 6th edition*. Thousand Oaks, CA: Sage Publication.

Rossi, P.H., Lipsey, M., and Freeman.H.E. (2004). *Evaluation: A Systematic Approach 7th edition*. Thousand Oaks, CA: Sage Publication.

Rossi, P.H., Berk, R.A., and Lenihan, K.J., (1980). *Money, Work and Crime: Some Experimental Evidence*. New York: Academic Press.

Tan,J.P.,J. Lane, and G.Lassibille (1999) *Schooling Outcomes in Philippine Elementary Schools: Evaluation of the Impact of Four Experiments*, In

World Bank Economic Review, September.1999.
USAID (1987). *Rapid, Low-Cost Data Collection
Methods for A.I.D.* Washington D.C.: USAID.
Valadez, J., and Bamberger, M. (1994). *Monitoring
and Evaluating Social Programs in Developing
Countries: A Handbook for Policymakers,*

Managers, and Researchers. Washington, D.C.:
World Bank.
佐々木亮 (著) 龍慶昭 (監修) (2003) 『政策評
価トレーニングブック』、多賀出版
(2006.2.10受理)

表3 Poverty Action Labの登録案件一覧

プロジェクト名; 報告書名 Project name; Title of a research paper	サンプル Sample	対象国 Country	分野 Sector	評価費用負担者 Funding
アジア				
Women as policy makers; Evidence from an India-Wide Randomized Policy	Village councils	インド	分権化、ジェンダー	National Institute of Health, MacArthur Foundation
Balsakhi Program; Evidence from Two Randomized Experiments in India	Primary Schools	インド	初等教育	World Bank, ICICI corp., MacArthur Foundation
Balwadi Health Program; Iron Deficiency Anemia and School Participation	4,000children (2-6 yrs old)	インド	初等教育 / 保健	World Bank, UC Berkeley Center for Health Research, ICICI
Computer-Assisted Learning Project; (Paper Forthcoming)	100 primary schools	インド	初等教育	World Bank, ICICI Corp. MacArthur Foundation, etc.
Improving Teacher Attendance in Rural India; Holding Teachers Accountable: Evidence From a Randomized Evaluation in India	Non-formal Education Centers	インド	教育	MacArthur Foundation
Affirmative Action; (Paper Forthcoming)	College-age cohorts	インド	差別、教育	MIT and University of Chicago
Information and Community Mobilization; (Paper Forthcoming)	280 villages	インド	教育、参加型開発	World Bank
Healthcare and Health Status in Udaipur districts, Rajasthan; (Paper Forthcoming)	1,000 households	インド	保健	MacArthur Foundation, Center for Health and Well being at Princeton University, NIH, MIT
Commitment Savings Products; (Paper Forthcoming)	835 individuals	フィリピン	マイクロクレジット	ADB, Russell Sage Foundation, National Science Foundation
Group Versus Individual Liability; (Paper Forthcoming)	Micro-lending NGOs	フィリピン	マイクロクレジット	World Bank
Combating Corruption in Community Development; Monitoring Corruption: Evidence from a Field Experiment in Indonesia	600 villages	インドネシア	地方開発	DFID-World Bank Strategic Poverty Partnership Trust Funds
Political Institutions and Local Public Goods; (Paper Forthcoming)	500 villages	インドネシア	地方開発	Unavailable
アフリカ				
Measuring the Impact of School Inputs Kenya; Retrospective vs. Prospective Analysis of School Inputs: The Case of Flip Charts in Kenya	178 Primary Schools*	ケニア	初等教育	World Bank Research Committee, National Science Foundation

Primary School Deworming Project; Worms: Identify Impacts on Health and Educatin in the Presence of Treatment Externalities	75 primary schools	ケニア	初等教育 / 保健	World Bank, Partnership for Child Development, National Institute of Health
Incentives to Learn; Incentives to Learn	128 primary schools	ケニア (Busia)	初等教育	World Bank, Young Green Foundation, MacArthur Foundation
Teacher Incentives; Teacher Incentives	50 primary schools *	ケニア	初等教育	World Bank, MacArthur Foundation
Decentralization: A Cautionary Tale; Decentralization: A Cautionary Tale	14 primary schools *	ケニア	初等教育	World Bank
The Rockefeller Effect; The Rockefeller Effect	80 women's groups *	ケニア	ジェンダー	World Bank Social Capital Initiative, National Bureau of Economic Research (NBER)
Textbook and Test Scores; Textbooks and Test Scores: Evidence from a Prospective Evaluation in Kenya	100 primary schools *	ケニア	初等教育	World Bank Research Committee, Science Foundation
The Illusion of Sustainability; The Illusion of Sustainability	75 primary schools *	ケニア	初等教育	World Bank, National Institute of Health Fogarty International Center, UC Berkeley Center Health Research
Social Learning about Health in Kenya	75 primary schools	ケニア	初等教育	World Bank, National Institute of Health Fogarty International Center, UC Berkeley Center Health Research
HIV/AIDS Prevention Education in Primary Schools; (Paper forthcoming)	360 primary school *	ケニア	保健	World Bank
Finding Missing Markets: An Agricultural Brokerage Intervention in Kenya; (Paper Forthcoming)	100 self-help groups	ケニア	マイクロクレジット、 地方開発	World Bank, IDRC, SAGA
Understanding Technology Adoption: Fertilizer in Western Kenya; (Paper Forthcoming)	Kenyan Farmers	ケニア	地方開発	MacArthur Foundation
Kenya Life Panel Survey; (Paper Forthcoming)	7,500 young adults	ケニア	教育、保健	World Bank, National Institute of Health Fogarty Int'l Center, UC Berkeley Center Health Research
School Meals; (Paper Forthcoming)	50 primary schools *	ケニア	初等教育、保健	MacArthur Foundation, Colin-Wauters Fund, Flemish Science Fund, & Univ. of Antwerp
Extra-Teacher Provision (ETP) ; (Paper Forthcoming)	210 primary schools	ケニア	初等教育、分権化	World Bank
Interest Rates and Consumer Credit; (Paper Forthcoming)	Consumers	ケニア 南アフリカ	マイクロクレジット	Unavailable

Marketing Effects in a Consumer Credit Market; (Paper Forthcoming); (Paper Forthcoming)	Bank customers	南アフリカ	マイクロクレジット	Unavailable
Improving Primary Education Management in Madagascar (AGEMAD); (Paper Forthcoming)	Primar	マダガスカル	初等教育	World Bank
中南米				
Business and Health Education for Microfinance Clients; (Paper Forthcoming)	Lending groups	ペルー	マイクロクレジット、保健	World Bank, Henry E. Niles Foundation
Valuing Trust in Shantytowns; (Paper Forthcoming)	30 neighborhoods	ペルー	地方開発	Russell Sage Foundation
School Choice in Colombia: Vouchers for Private Schooling Colombia: Evidence from a Randomized Natural Experiments	1,600 applicants	コロンビア	初等教育	World Bank, National Institute of Health, National Science Foundation
Long-Term Impacts of School Choice in Colombia; Long-term Consequences of Secondary School Vouchers: Evidence from Administrative Records in Colombia	1,600 applicants	コロンビア	初等教育	National Science Foundation
北米				
Discrimination in the Job Market: Are Emily and Greg More Employable than Lakisha and Jamal?	Chicago and Boston Firms	米国	差別、ジェンダー	Unavailable
The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment	University staff employees	米国	マイクロクレジット	National Science Foundation
Peer Effects, Alcohol, and College Roommates; Peer Effects and Alcohol Use Among College Students	Undergraduate students	米国	高等教育	MacArthur Foundation, Joint Center for Poverty Research
Peer Effect, Diversity, and College Roommates; Empathy or Antipathy? The Impact of Diversity	Undergraduate students	米国	高等教育	MacArthur Foundation, Joint Center for Poverty Research
Saving Incentives for Low and Middle Income Families: Evidence from a Field Experiments with H & R Block	60 H&R Block offices	米国	貯蓄	H&R Block, Retirement and Security Project, Brookings
Time Management; (Paper Forthcoming)	Respondents	米国	行動経済調査	Russell Sage Foundation
Discrimination in the Judicial System; (Paper Forthcoming)	Chicago, Boston, & Federal Judges	米国	差別、ジェンダー	MIT and University of Chicago
Bank Accounts: Take Up and Impact; (Paper Forthcoming)	Poor people	米国	マイクロクレジット	Unavailable

(注1) ケニアで実施された案件の多くが、Busia District, Teso Districtの両県を対象地域としている。サンプルの欄に*の記載のある案件がそれぞれに該当するが、同一案件内での実施か、それぞれ個別の案件かは不明である。

(注2) セクター(Sector)は、Poverty Action Labの分類を参考にして、著者が再分類して記載した。
(source) Poverty Action Lab(<http://www.povertyactionlab.com/projects>)

Challenge of Evidence-Based Evaluation in ODA Sector: Poverty Action Lab

Ryo Sasaki

Western Michigan University
Ryoh.sasaki@wmich.edu

Abstract

A challenge of evidence-based evaluation has already started at official development assistance (ODA) sector, just like the trial of Cochran Collaboration in Medicine and that of Campbell Collaboration in social policies. The name of the project is MIT's Poverty Action Lab. It has actively conducted and produced evaluation results using rigorous randomized experimental design since its establishment in 2003. In this paper, firstly the progress of the project is reviewed, and then, based on its review, several issues relating to randomized design are discussed. They include ethical issue, cost issue, bias issue, the reason why true experimental design is more preferable than quasi-experimental design, and the reason why this design has not been used in this sector. Then, the current situation and issues in Japanese ODA relating to introduction of experimental design is discussed. Finally it is recommended that Japan should actively commit this project because Japan is the largest financier of the World Bank that has financially supported this project.

Keywords

experimental design, ODA, request-based, Poverty Action Lab

【研究論文】

都市大気汚染政策における社会的能力の評価

村上 一真 松岡 俊二

広島大学大学院国際協力研究科

mkazuma@hiroshima-u.ac.jp

smatsu@hiroshima-u.ac.jp

要 約

本研究は、能力の評価基準を実証的に示し、それを能力評価に係るフレームワークに位置づけることで、新たな能力評価手法を提案した。まず、日本の都市の大気汚染政策を対象に、行政の大気汚染政策に係る能力要素として、「科学的知見」の提供能力、「環境政策資源」の運用能力、「直接規制」の遂行能力、「経済的支援」の遂行能力の4つを、因子分析により抽出した。その上で、この能力要素の大気質改善への貢献とウェイトを実証的に確認し、評価すべき能力要素、すなわち、能力の評価基準として示した。以上を踏まえて、政策サイクルにおける能力要素の役割を設定し、能力評価に係るフレームワークを示した。このフレームワークであるアクター・ファクター・マトリクスを用いたアクター・ファクター分析は、UNDPで示されている能力評価手法と比較して、評価基準の体系的な整理、効率のかつ効果的な評価実施、能力形成に向けた適切な政策提案の3点において優位性が指摘できた。

キーワード

Capacity Development、評価手法、能力指標、因子分析、大気汚染対策

1. 研究の背景と目的

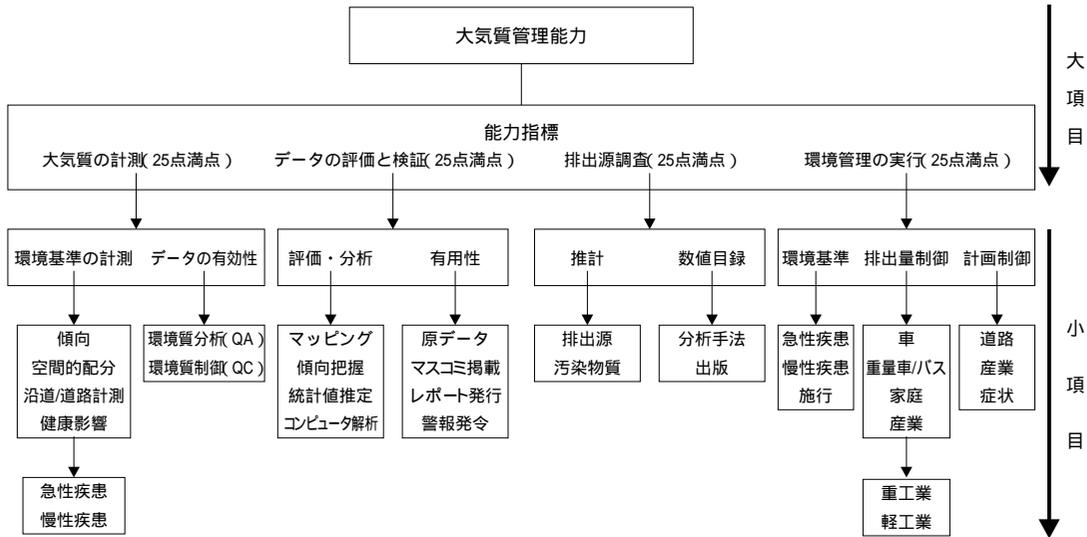
開発援助におけるCapacity Developmentの概念は、援助の質向上に向けたキーワードとして、1990年代初頭より多く用いられるようになり、UNDP、CIDA、JICA等の援助機関で、その実践に向けた議論が進められてきた（UNDP 1998、Lavergne and Saxby 2001、JICA 2004）。そして、貧困に起因する環境破壊への環境ODA増加や、地球環境問題の顕在化に対し、環境問題に対処する能力への着眼も必然であった。

環境問題に対処する能力は、1987年のブルント

ラント委員会報告による持続可能な発展提唱の流れを受け、1992年の国連環境開発会議で採択されたアジェンダ21で必要性が示された¹⁾。そして、それを受けたOECD/DACで、Capacity in Environment、Capacity Development in Environmentの理念や定義の議論²⁾がなされてきた（OECD/DAC 1999）。

しかし、能力への着目と能力開発の必要性の認識は共有されながらも、いずれも、能力とは何かについて、統一的な合意が得られたわけではなく、抽象的、概念的に定義するに留まっている。これに関して、松岡・本田（2002）は、「概念が広く

図1 都市の大気質管理能力の内部構造 (UNEP/WHO 1996)



(出所) 松岡他 (2000)

抽象的であるが故、理念として援助機関に受け入れられてきたが、定義や具体的な中身について十分な検討は行われていない」と指摘している。また、Boesen et al. (1998)、Janicke et al. (1997)、UNDP/GEF (2003) など、具体的に能力の構成要素を示している場合でも、それを論理的および体系的に示すには至っていない。

これらに対し、UNEP/WHO (1996) は、都市の大気質管理に係る能力の内部構造を定量的に明らかにしようとした (図1)。ここでは、経験的見地より、大気質の計測、データの評価と検証、排出源調査、環境管理の実行の4項目を能力の構成要素として挙げ、20都市の大気質管理に係る能力の得点化と評価を行っている。しかし、能力要素の必要十分性は検証されず、能力要素ごとの大気質改善への寄与度も示されておらず、評価対象として設定された能力要素の妥当性に疑問が残る。また、各能力要素は、下位の評価項目数や大気質改善に向けた役割の違いにも関わらず、25点満点で重みづけなしに並列に評価されるなど、能力概念の明確化とその測定方法について十分な研究蓄積があるとはいえない。

能力開発には、事前の能力評価が必要となる。しかし、対象とする能力が明確にされ、適切な評

価基準が設定されなければ、効果的な評価は行えない。環境問題に対処する能力の評価手法確立のためには、まず、評価すべき能力とは何かについて、理論的かつ実証的な観点から、体系的に示す必要がある。ここでの評価すべき能力とは、政策成果に貢献する能力であり、これを評価基準として設定することが、効率的で効果的な能力評価手法の提案につながる。

本研究は、環境質改善に貢献する能力要素を実証的に示し、それを評価基準とする実践的な能力評価手法の開発を目的とする。具体的には、環境対策に係る能力要素を抽出し、政策成果である環境質との関係性を実証的に示すことで、環境質改善に貢献する能力要素を提示する。そして、これを評価基準とする新たな能力評価手法を、松岡他 (2004) の「社会的環境管理能力」を雛形にして提案し、UNDP (1998) での既存の能力評価手法との比較を通じて、その優位性を明らかにする。

本研究の構成は、2.で研究対象および方法として、実証分析での分析対象、分析手法およびデータを示し、3.で分析結果を示す。4.では、3.で明らかにした成果をもとに、能力評価手法の提案を行い、5.で結論を述べる。

2. 研究対象および方法

(1) 分析対象の設定

UNEP/WHO (1996) との比較を想定し、環境対策分野として、都市の大気汚染対策を対象とする。大気汚染物質は、産業型公害の典型である二酸化硫黄 (SO₂) とする。分析期間は、全国の長期継続測定局のSO₂濃度年平均値が1967年をピークに減少に転じていることを踏まえ、原嶋・森田 (1995) での環境政策の前進期 (1965~)、松岡・朽木 (2003) での社会的環境管理能力の本格的稼働期 (1970~) という、環境政策が本格的に実施され、成果を挙げ始めた時期からを対象とする。ここでは、データの制約上、1970年から2000年までとする。

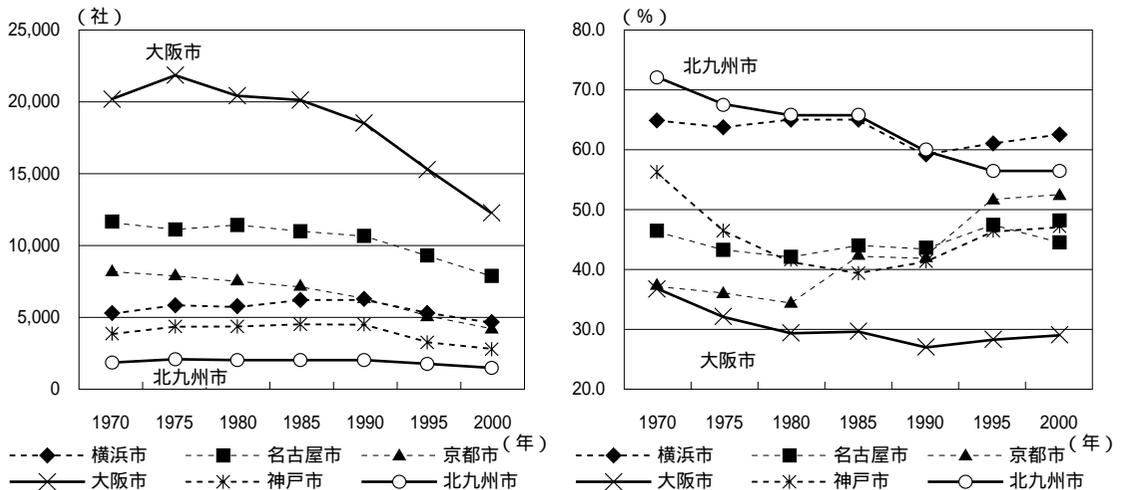
分析対象とする能力は、行政、企業、市民の各主体の能力が候補に挙がる。これは、松岡他 (2004) の社会的環境管理能力の定義である「政府・企業・市民の3つのアクター (主体) およびアクター間の相互関係からなる、環境問題に対処するための社会全体としての総体的な能力」から導かれる。ここで、産業型公害としてのSO₂は、市場の失敗としての外部不経済に起因するものであり、行政介入の根拠に基づき、環境政策、特に

直接規制が対策に係る主要な役割を担った (植田1996)。具体的には、行政が汚染排出者である企業に対して、一定の権限を持って、燃料転換、運転管理、技術開発・導入などの汚染抑制行動を促すことで、大気質改善が進められた。また、1970年前後は、いわゆる公害国会での公害関係14法案の制定・改定 (1970年)、環境庁設置 (1971年) など、環境政策が充実強化された時期である (原嶋・森田 1995)。したがって、本研究では、分析対象とする能力を、行政の能力とする。

対象とする都市は、政令指定都市であり、都市構造の異なる北九州市と大阪市の2都市とする³⁾。図2のとおり、製造業事業所数および大規模事業所比率 (従業員300人以上事業所の全製造品出荷額に占める割合) の違いより、2都市の都市構造は大きく異なる。1970年時点での全政令指定都市の比較において、北九州市は製造業事業所数は最も少なく、大規模事業所比率は最も高い。一方、大阪市は製造業事業所数は最も多く、大規模事業所比率は最も低い。また、行政の財政規模、職員数は大阪市が最も多く、北九州市は京都市と並んで最も少ない。そして、人口、人口密度は大阪市が最も高く、北九州市は最も低い。

これらの産業構造を中心とした都市構造の違いは、工場・事業所などの固定排出源からの排出が

図2 製造業事業所数、および従業員300人以上事業所の全製造品出荷額に占める割合



(出所) 経済産業省「工業統計表」(各年版)より筆者作成

表1 都市構造に起因する大気汚染政策（SO₂）の違い

	北九州市	大阪市
主な規制対象	環境対策に向けた技術力、資金力の豊富な少数の大企業	環境対策に向けた技術力、資金力に余裕が無い多数の中小企業
政策手段	公害防止協定	行政指導
政策内容	1970年施行の北九州市公害防止条例の規定に基づき、1972年に「いおう酸化物に係る公害の防止に関する協定」（第一次）が、北九州市、福岡県と47社54工場の間で一括締結された。 公害防止協定は、行政と企業が結ぶ紳士協定であり、法律以上の上乘せ規制、未規制事項の措置等があり、契約文書の形態で調印された。事業所別の目標値などを記した別紙計画書は、非公表となっている。	科学的根拠を持った「行政指導計画」を公表し、工場・事業所に対しての説明会で、大気汚染状況や対策技術・支援制度などを説明し、大阪市の技術職員が立入調査を行い、個別の工場・事業所に具体の改善目標を示した。そして、双方で議論を重ね、各企業に「大気汚染防止計画書」を年次計画として市長宛に提出させ、その実行を求めた。 大気汚染防止計画書には、市長名の承認証書は交付されず、企業の自主的な提出という形での対策実施となり、公文書のない口頭での行政指導が特徴といえる。
コミュニケーション	1970年に北九州市、福岡県、福岡通商産業局と市内企業（30社32工場）から構成される「北九州市大気汚染防止連絡協議会」が設立され、発生源規制等に関する事前協議や情報交換が行われた。 この企業会員30社32工場は、当時、硫黄酸化物排出量の市内全排出量の97%を占めており、事前協議の結果は、施策の実現性と実効性の確保につながった。	大阪市が指導し行政区ごとに設立した「区煤煙防止会」を通じて、燃焼指導、技術・知識の普及を図った。 また、煤煙防止に関する調査研究や普及啓発活動を行う「大阪市煤煙防止会連合会」の協力を得て、全市域を対象とした業種別説明会、施設規模別説明会など、各種説明会も随時開催された。
科学的調査と目標設定	1969年から実施された北九州地区産業公害総合事前調査による汚染状況の把握と汚染予測などの科学的調査の結果に基づき、汚染削減の目標値を設定した。そして、企業に自社の汚染状況の把握と対応方針が記載された「硫黄酸化物に係る改善計画書」の提出を求め、その進捗確認のため、行政による立入り検査による監視が行われた。	汚染の著しい西淀川区においては、西淀川区公害特別機動隊や大気汚染解析班を発足させ、夜間を含む立ち入り調査、全工場へのヒアリングによる詳細な現状把握によりきめ細かい指導を行った。また、大気拡散シミュレーションにより、各工場ごとの汚染寄与率の算定を行い、企業ごとの公害防止計画の目標値を設定した。
技術開発・導入	八幡製鉄所がクリーナー・プロダクション化を図り、省エネ、硫黄酸化物排出量低減を果たすなど、企業自らが技術力、資金力を背景に技術開発・導入を行うことが可能だった。また、中小企業も多くは大企業の下請けで、大企業の敷地内に立地している場合も多く、大企業に対する指導を通じて、中小企業も間接的に指導できた。	西淀川区での企業の防止装置の導入に際して、「大阪市公害防止設備資金融資制度」が大いに活用された。また、「大阪市公害工場移転跡地買収制度」は独自の経済支援制度として特筆すべきと評価されているように、行政が対策技術や資金調達を支援していた。 また、大阪役所の職員は頻繁に対策技術の見本市に通い、企業への説明会には対策装置のパフレットを持参して臨んだとされる。

（出所）大阪市（1994）、北九州市（1998）、勝原（2001）、藤倉（2002）より筆者作成

主なSO₂に対して、汚染排出状況と対策の相違を生じさせる。表1にあるように、北九州市の主な規制対象は、環境対策に向けた技術力、資金力の豊富な少数の大企業であり、行政と企業のコミュニケーションに係る取引費用は小さく、定期的な協議の場で行われる事前協議や情報交換により、政策の実現性と実効性が担保できた。他方、大阪市は、環境対策に向けた技術力、資金力に余裕が無い多数の中小企業が主な対象であり、コミュニケーションに係る取引費用は大きく、個別対応の積み重ねにより政策の実現性と実効性の確保を目指すほかなかった。これらより、両市のSO₂政策は、企業の対策促進を促すための取引費用の大きさが要因の一つとなり、北九州市が公害防止協定、大阪市が行政指導という異なる政策手段選択として帰結する⁵。

したがって、都市規模を踏まえた合理的な政策手段を用いた典型例として、北九州市と大阪市を選択し、都市規模および政策手段の異なる都市ごとに大気汚染政策に係る能力の構成要素および大気質改善への貢献を明らかにする。そして、それらを比較検討することで、都市ごとおよび共通的な結論を導き出す。

（2）分析手法

大気汚染政策に係る能力要素を抽出し、それらが大気質改善に貢献する能力要素かを確認するため、2段階の実証分析を行う。

分析1：能力要素の抽出

行政の大気汚染政策に関するデータを用いて因

因子分析を行い、能力要素を抽出する。因子分析は、いくつかの顕在化した事象の背後にある、共通で潜在的な変量を抽出する分析手法である。ここでは、いくつかの大気汚染政策が、共通的な因子ごとに区分される。この区分された複数の大気汚染政策の潜在的な共通性を、それら政策の実効性を担う能力として捉え、因子を行政の大気汚染政策に係る能力要素として提示する。因子分析は、能力という潜在的で抽象的な概念を可視化させ、また、因子ごとの能力全体に占める寄与率も定量的に把握できるため、本研究の目的に沿った分析手法となる。

分析2：能力要素の検証

分析1で示された各能力要素、およびウェイト付けされた能力要素により統合された大気汚染政策に係る能力全体が、大気質改善に寄与しているかを検証する。つまり、抽出された能力要素に関して、個および全体としての政策成果への貢献の有無を検証する。このことは、各能力要素の存在とウェイトの妥当性の検証にもなる。

各能力要素の検証は、分析1の因子分析により得られた能力要素ごとの因子得点を説明変数とし、SO₂濃度を目的変数とした回帰分析を行い、各能力要素と大気質の関係性を明らかにする。

各能力要素を統合した能力の検証は、まず、因子分析により得られた能力要素ごとの因子得点を、因子負荷量の寄与率をウェイトとして加重平均し、大気汚染政策に係る統合された能力を導出する⁶。そして、それを説明変数とし、SO₂濃度を目的変数として回帰分析を行い、大気汚染政策に

係る能力全体と大気質の関係性を明らかにする。因子負荷量の寄与率は、大気汚染政策に係る能力全体を各因子がどの程度説明できるかを示しており、能力要素ごとのウェイトとして捉えることができる。したがって、ウェイトづけされた能力要素により導出された大気汚染政策に係る能力と、政策成果である大気質との回帰分析において、適切な関係性が確認されることで、大気汚染政策に係る能力の妥当性、すなわち、各能力要素のウェイトの検証がなされる。

(3) データ

分析対象とする行政の範囲は、市に加えて、政策形成・遂行に向けて科学的データ・情報提供を行った、公害対策に係る市立の研究所を対象とする⁷。ここでは、北九州市環境科学研究所（旧北九州衛生研究所）、大阪市環境科学研究所（旧市立大阪衛生試験所）が対象となる。

データは、両市ともに1970～2000年において継続的に収集可能な、各組織の直接的な大気汚染政策に関わるインプット、アウトプットに関する量的データとする（表2）。SO₂濃度改善という政策成果（アウトカム）との比較に際して、その達成に向けた大気汚染政策に関する行政活動の程度を示すデータを対象とすることとなる。

研究所では、研究所予算、研究所人材（在籍人数×平均勤務年数）、研究者の学会発表数、研究者の学術雑誌への論文掲載数、研究所による行政や住民からの依頼に基づく大気汚染に係る検体の分析件数が収集できた。

表2 データ

データ	出所
環境科学研究所・予算	
環境科学研究所・人材（人数×平均勤務年数）	北九州市環境科学研究所報（各年版）
環境科学研究所・学会発表数	大阪市立環境科学研究所報告（各年版）
環境科学研究所・論文掲載数	地方公務員給与制度研究会「地方公務員給与の実態」（各年版）
環境科学研究所・大気汚染検査件数	
市・環境部署人材（人数×平均勤務年数）	北九州市の環境（各年版）
市・工場事業場への大気汚染立入検査施設数	大阪市環境白書（各年版）
市・企業への大気汚染防止資金融資額	地方公務員給与制度研究会「地方公務員給与の実態」（各年版）
市・企業への大気汚染防止資金融資件数	
市・衛生費	地方財政調査研究会「地方財政統計年報」（各年版）
一般環境大気測定局数	環境省「日本の大気汚染状況」（各年版）

（出所）筆者作成

市では、衛生費⁸、一般環境大気測定局数、環境部署人材（在籍人数×平均勤務年数）、大気汚染防止法や公害防止条例および市民からの苦情等に基づく工場・事業場への大気汚染に関する立入検査件数、要綱に基づく企業への大気汚染防止資金融資件数、およびその融資額が収集できた。

ここで、人材は、量だけでなく、質の把握が重要となる（Honadle 2001）。Barro（1991）、Mulligan et al.（1997）ら、経済成長に係る人的資本の実証分析では、いわゆる労働力の質の計測において、主に教育年数や労働者所得などが代理変数として用いられている。これらを踏まえて、環境政策に携わる人材の質を、経験および熟練を示すものとして平均勤務年数を設定し、人員数と掛

け合わせて人材として示した。

3. 分析結果

（1）分析1：能力要素の抽出

探索的因子分析（主因子法・プロマックス回転）により、スクリーテストを基準として2市でそれぞれ4因子が抽出された（表3、表4）。クロンバックの係数は、大阪市の第4因子が0.684と若干低くなったが、その他は0.782～0.976となり、各因子の存在について、一定の信頼性が確保できた。因子ごとに含まれるデータ構成は、1つのデータ

表3 因子負荷量と能力要素抽出（北九州市）

データ	因子1	因子2	因子3	因子4
環境科学研究所・予算	0.933	-0.182	0.000	0.058
北九州市・衛生費	0.819	-0.080	0.380	0.342
環境科学研究所・人材（人数×平均勤務年数）	0.733	0.310	0.411	0.347
一般環境大気測定局数	0.692	0.172	0.502	0.408
北九州市・環境部署人材（人数×平均勤務年数）	0.096	0.915	-0.076	0.216
北九州市・工場事業場への大気汚染立入検査施設数	-0.229	0.855	0.167	0.024
環境科学研究所・大気汚染検査件数	0.133	0.707	-0.033	0.450
北九州市・企業への大気汚染防止資金融資額	-0.198	0.073	-0.818	-0.100
北九州市・企業への大気汚染防止資金融資件数	-0.571	-0.372	-0.603	-0.286
環境科学研究所・学会発表数	0.394	0.253	0.170	0.864
環境科学研究所・論文掲載数	0.193	0.420	0.271	0.526
固有値	3.363	2.508	1.821	1.594
寄与率（%）	52.0	21.4	6.9	4.2
累積寄与率（%）	52.0	73.4	80.2	84.4

（出所）筆者作成

表4 因子負荷量と能力要素抽出（大阪市）

データ	因子1	因子2	因子3	因子4
一般環境大気測定局数	0.971	0.189	-0.034	0.002
大阪市・環境部署人材（人数×平均勤務年数）	0.832	0.443	0.220	0.216
環境科学研究所・人材（人数×平均勤務年数）	0.687	0.629	0.225	0.213
環境科学研究所・予算	0.665	0.604	0.257	0.317
大阪市・衛生費	0.613	0.542	0.381	0.409
大阪市・企業への大気汚染防止資金融資額	-0.225	-0.952	-0.046	-0.049
大阪市・企業への大気汚染防止資金融資件数	-0.492	-0.827	-0.052	-0.108
環境科学研究所・論文掲載数	0.068	-0.029	0.992	-0.091
環境科学研究所・学会発表数	0.389	0.496	0.580	0.230
大阪市・工場事業場への大気汚染立入検査施設数	0.020	-0.383	-0.503	-0.568
環境科学研究所・大気汚染検査件数	-0.212	-0.103	0.043	-0.489
固有値	3.538	3.479	1.995	0.891
寄与率（%）	42.6	30.3	10.7	4.9
累積寄与率（%）	42.6	72.9	83.5	88.4

（出所）筆者作成

を除いて2市で同じ結果となり、2市とも同様の因子の解釈が可能となった。

北九州市は、第1因子はヒト、モノ、カネなど、環境政策を資源面で支える「環境政策資源」の運用能力、第2因子は汚染排出源への立入りや検体検査としての「直接規制」の遂行能力、第3因子は汚染排出者の対策を経済的に支援する「経済的支援」の遂行能力、第4因子は政策形成・遂行の根拠となる科学的根拠を示す「科学的知見」の提供能力と解釈した。一方、大阪市は、第1因子は「環境政策資源」の運用能力、第2因子は「経済的支援」の遂行能力、第3因子は「科学的知見」の提供能力、第4因子は「直接規制」の遂行能力となった。

両市を比較すると、「市・環境部署人材」データが、北九州市では「直接規制」の遂行能力、大阪市では「環境政策資源」の運用能力となった以外は、因子の序列は異なるが、全く同じ因子構造となった。この「市の環境部署・人材」での相違は、両市の大気汚染政策の違いで述べたように、北九州市では少数の企業を対象とした公害防止協定に基づく立入検査で主に人材の持つ機能が発揮され、大阪市では多数の企業を対象とした行政指導において、数多くの説明会実施や技術調査など、広範な行政活動において人材の持つ機能が発揮されたという理解ができる。

また、因子ごとの寄与率は、第1因子の「環境政策資源」の運用能力が、北九州市では52.0%、大阪市では42.6%となり、大気汚染政策に係る能力の約5割程度を説明する。そして、第2因子を合わせると、能力の7割強が説明できる。第2因子は、北九州市は「直接規制」の遂行能力、大阪市では「経済的支援」の遂行能力となった。これは、環境対策に向けた技術力、資金力の豊富な少数の大企業を対象とした北九州市では、企業の対策の事後的な監視が主な活動となり、環境対策に向けた技術力、資金力に余裕が無い多数の中小企業を対象とした大阪市では、企業の対策過程での指導・支援が主な活動となるという、都市構造の違いに基づく政策手段の違いで説明できる。両市とも、各都市の特性に適した政策を遂行する能力要素と、その政策形成等を可能にする資源を運用する能力要素により、能力の大部分が説明できること

が示された。

異なる都市形態で、異なる政策手段を用いたにも関わらず、両市とも行政の大気汚染政策に係る能力は、累積寄与率の大きさを踏まえると、この4要素で説明できることが実証的に示された。そして、能力要素の第1、2因子の寄与率の大きさより、両市の大気汚染政策の実際と整合的であることも確認された。

(2) 分析2：能力要素の検証

SO₂濃度を目的変数とし、4つの能力要素の因子得点を説明変数として、最小2乗法（OLS）による回帰分析を行った。両市とも全ての係数の符号が負で、かつ統計的に有意となり、4つ全ての能力要素のSO₂濃度改善への寄与が明らかとなった（表5）。また、標準偏回帰係数より、両市とも「環境政策資源」の運用能力が最も大きい寄与を示した。次いで、両市ともに、「経済的支援」の遂行能力、「直接規制」の遂行能力、「科学的知見」の提供能力と続いた。

表3、4では、大気汚染政策の実際を反映する形で、能力全体に占める能力要素の寄与率の序列は両市で異なったが、表5では、能力要素のSO₂濃度改善へ寄与度の序列は、両市とも等しくなった。表5より、政策成果の達成に向けた能力形成の優先度は、「環境政策資源」、「経済的支援」、「直接規制」、「科学的知見」とするのが合理的となる。ここで、両市とも「環境政策資源」の能力全体に占める寄与率は最も大きく、能力形成の優先度と

表5 4つの能力要素の回帰分析

	北九州市	大阪市
科学的知見	-0.366 ** (-6.49)	-0.113 * (-2.51)
環境政策資源	-0.567 ** (-10.01)	-0.881 ** (-19.55)
直接規制	-0.368 ** (-6.54)	-0.222 ** (-4.94)
経済的支援	-0.501 ** (-8.85)	-0.323 ** (-7.21)
定数項	0.010 ** (18.06)	0.016 ** (26.79)
F 値	72.587 **	118.367 **
Adj.R2	0.905	0.940

** 1%有意、* 5%有意

(出所)筆者作成

表6 大気汚染政策に係る能力の回帰分析

	北九州市	大阪市
大気汚染政策に係る能力	-0.807 ** (-7.36)	-0.935 ** (-14.19)
定数項	0.010 ** (9.26)	0.016 ** (18.19)
F 値	54.155 **	201.247 **
Adj.R2	0.639	0.870

** 1%有意、* 5%有意

(出所) 筆者作成

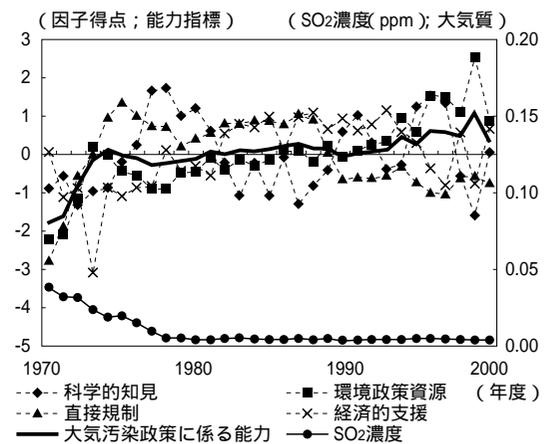
して適切であった。しかし、「経済的支援」は、大阪市では、寄与率が2番目に大きい第2因子となったのに対し、北九州市では第3因子となっており、大阪市がSO₂濃度改善に向けて、より合理的な対応を行ったといえる。

次に、4つの能力要素の因子得点を、因子負荷量の寄与率をウェイトとして加重平均した値を算出し、それを大気汚染政策に係る能力として設定した。そして、それを説明変数とし、SO₂濃度を目的変数として、最小2乗法による回帰分析を行った(表6)。結果、両市とも係数が負符合で、かつ統計的に有意となり、大気汚染政策に係る能力が大気質の改善に寄与することが確認された。

さらに、1970~2000年における大気汚染政策に係る能力の形成過程を図示した(図3、図4)。両市とも1970年代前半に大気汚染政策に関する能力形成が急激に進んだことが見て取れる。原嶋・森田(1995)は、日本の環境政策の発展段階として、環境政策が充実強化されその効果が認められた前進期(1965~1974年)、環境政策の対象となる環境問題が質的に変化する変容期・停滞期・地球環境時代(1975年~)に区分している。また、寺西(1994)は日本の環境政策の歴史的推移を、形成・発展期(1965~1974年)、激流・後退期(1975~1988年)、模索・再編期(1989年~)に区分している。

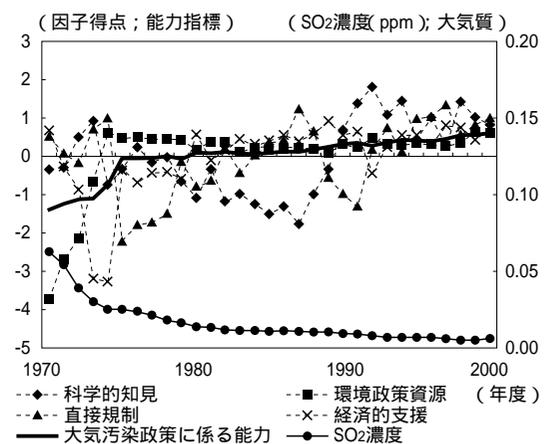
本分析の大気汚染政策に係る能力形成過程は、これらの環境政策区分と整合的であり、かつ、政策成果である大気質の改善にも貢献している。このことは、環境政策の実効性を担う能力も同様の経緯で発展し、政策成果を挙げてきたことを示している。政策成果である大気質と政策の実効性を担う能力の関係性が確認されたことから、大気汚染政策に係る能力の妥当性、すなわち、4つの能

図3 能力形成と大気質の推移(北九州市)



(出所) 筆者作成

図4 能力形成と大気質の推移(大阪市)



(出所) 筆者作成

力要素の存在とウェイトが検証された。

4. 大気汚染対策に係る能力評価手法の提案

3.では、因子分析により4つの能力要素を抽出し、それらがいずれも大気汚染改善という政策成果に貢献する機能を有すること、また、能力全体に占める各能力要素のウェイトは妥当であることを実証的に確認した。これにより、4つの能力要素が評価すべき能力要素、つまり、能力の評価基

準として適当であることが示された。

新たな能力評価手法の提案に際しては、これらの能力要素を、能力評価に係るフレームワークの中に位置づける必要がある。

ここでは、政策の実効性を担う能力の水準が、政策成果の程度を規定すると捉え、4つの能力要素が政策サイクルにおいてそれぞれ担う役割を示す。能力要素の関係性を、政策サイクルにおける役割の関係性として提示することで、能力の評価基準を政策サイクルにおける各ステージの評価基準として同時に示すことができる。このことは、能力評価に係るフレームワークが、多面的な評価の視点を内包することとなる。

また、政策サイクルは知識マネジメントの観点から捉えることができ（山内他 2000）、政策サイクルを知識サイクルとしても表現することで、能力要素の政策サイクルにおける役割設定の指針とした。

この能力要素の政策サイクル、知識サイクルとの対応は、能力要素間の関係性の明確化とともに、能力概念の具現化を手助けし、能力評価の際のデータ収集等を容易にさせる実用的なメリットがある。

(1) 能力評価に係るフレームワークの設定

政策成果の程度は、政策の実効性を担う能力水準により規定される。他方、政策成果は、政策課題設定、政策形成、政策遂行、政策評価のステージで構成される、一連の政策サイクルの円滑な循環（スパイラルアップ）により生み出されるとも説明できる。ここで、政策の実効性を担う4つの能力要素ごとに、政策サイクルでの役割を示すことが可能となる。もちろん、能力は特定の目的のみに対応する性質ではないため、能力要素が実効性を担う政策サイクルのステージは、明確に分離されるとは限らず、一部重複することもあり得る。

政策サイクルにおいて、4つの能力要素が果たす役割は、図5のように、政策課題設定（知識創造・蓄積）としての「科学的知見」の提供能力、政策遂行（知識実践）としての「直接規制」および「経済的支援」の遂行能力、政策形成（知識利用）および政策評価（知識再生産）としての「環

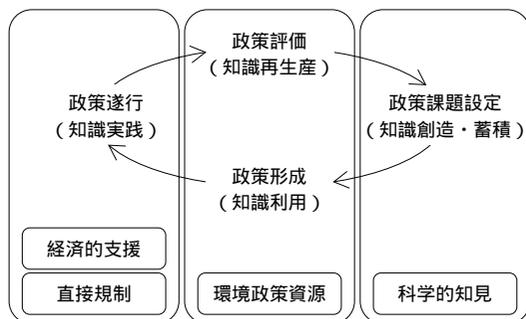
境政策資源」の運用能力と位置付けられる。また、後述するが、「環境政策資源」は、政策課題設定、政策遂行の実効性にも影響を与える。

「科学的知見」は、大気汚染状況の分析、汚染除去技術の評価等、調査研究、試験検査、情報収集に基づき、政策的に対応すべき課題を科学的根拠を基に示す役割を担う。勝原（2001）は、北九州市環境科学研究所を、科学的データ・情報を提供し説得力のある環境行政の実施を可能にしたとし、藤倉（2002）は、公害対策において、行政が企業に対策を迫る際の科学データの有効性を指摘し、両市の早期の研究所の設置を評価している¹⁰。

「直接規制」、「経済的支援」には政策遂行を担う役割が与えられる¹¹。「直接規制」は、大気汚染防止法や公害防止条例および市民からの苦情等に基づく工場・事業場への大気汚染に関する立入検査であり、燃料や原料採取による検査、機器の効率判定や維持管理状況検査などの届出内容や基準適合状況の検査とともに、公害防止組織の整備状況や公害防止管理者の活動状況の検査等を行う。そして、違反に対して、改善命令、改善勧告、改善指示などの行政指導を実施するなど、発生源対策を進める能力である。

「経済的支援」は、主に中小企業の公害対策を資金的に支援するもので、公害防止のために必要な機械、装置の設置、事務所や設備移転などに係る資金の融資、および利子助成を行う¹²。中小企業の規制基準達成の困難さに配慮して設計された制度であり、直接規制との相互関係に基づく、アメとムチの機能を有する大気汚染政策の一端を担

図5 能力評価に係るフレームワーク



(出所) 筆者作成

うものとして位置づけられる。「経済的支援」の執行能力は、単に融資枠の確保だけでなく、その過程での説明会実施や審査、管理監督などに係る能力として捉えられる¹³。

施設・設備、人材、予算などの「環境政策資源」の運用能力は、政策サイクルの全ステージにおいて、資源を確保し、適正に配分し、効率的に活用する役割を担う。施設・設備としての一般環境大気測定局は、環境政策の必要性および効果に関する情報を提供する。つまり、特に政策形成、政策評価のために、大気汚染状況を把握する資源といえる。

また、人材は環境政策に係る全ての行為の担い手であり、予算はその環境政策の対象範囲を定める枠組みとなる。人材、予算とともに、どのような政策を、どの程度実施するかを政治的な意思決定を司ることで、特に、政策形成および政策評価の機能を果たす。このことは、政策課題設定、政策遂行のステージでは、ある程度、目的と対応範囲が明確にされており、政治的な意思決定の必要性は相対的に低いと言える。

また、政策形成と政策評価を同様の役割に位置づけることは、政策評価論から説明できる。昨今の政策評価論では、政策アウトカムを想定し、政策の評価基準を定めた上で、政策形成を行うことが求められている（上山 1998、真山2001）。これは、“形成”された政策の成果を“評価”し、新たに政策を再検討するというスパイラルアップ型の政策サイクルにおいて、政策形成と政策評価は一連の政策項目、もしくは表裏の政策項目となっており、その実効性を支える能力も同様のものとして捉えられる。したがって、政治的意思決定を司る「環境政策資源」の運用能力が、政策形成と政策評価の役割をともに担うものとして説明できる。

以上、図5の政策サイクルにおける能力要素の役割の関係性、すなわち、実証分析により示された能力の評価基準を盛り込んだ能力評価に係るフレームワークが設定された。

(2) アクター・ファクター分析の提案

ここでは、図5の能力評価に係るフレームワー

クを用いた新たな能力評価手法を、松岡・朽木（2003）の「社会的環境管理能力」をもとに提案する。

実証分析では、政府という1アクターの大气汚染政策に係る能力要素を考察したが、この成果を応用し、企業、市民の評価項目の設定を行った。そして、政府・企業・市民という3つのアクターごとに、評価すべき能力要素（ファクター）を設定したアクター・ファクター分析を提案する。表7では、その分析フレームワークであるアクター・ファクター・マトリクスを示した。マトリクスでは、統計データにより定量的に把握できる評価細目を中心に示した。ファクターは、能力要素および政策サイクルでの3つの役割ごとに区分されており、3つのアクター区分とともに9つのセルを構成する。実際の評価では、このマトリクスを用いて、9つのセルで区分された大気汚染対策に係る能力を、統計データに加え、インタビューやアンケートでの質的情報を基に把握していく。

このアクター・ファクター分析は、体系的な能力要素の抽出、能力要素の大気質改善への貢献の確認、そして、政策サイクルに基づくフレームワークの設定を行った上で開発した能力評価手法であり、UNDP（1998）等で示されている個人、組織、社会システムという3区分での能力評価手法に比べて、以下の3点で優位性が指摘できる。

まず、評価基準の体系的な整理である。UNDP（1998）は、社会システムの能力要素について、政策、法・規制、マネジメント・アカウントビリティ、資源（人材、予算、情報）、プロセスを挙げている。そして、組織では、ミッション・戦略、組織文化・構造、能力（Competencies）、プロセス、資源（人材、予算、情報）、インフラを挙げている。しかし、これらは、能力の内部構造として、能力要素の必要十分性や関係性が明確にされておらず、体系的な評価基準設定となっていない。そのため、評価に際してのデータ収集・選択におけるフレームワークとしては不十分である。アクター・ファクター分析は、実証分析により体系的な能力抽出を行い、かつ、政策サイクルおよび知識サイクルに対応させる形で能力要素を設定している。このことは、政策サイクル、知識サイクル、能力要素という整合的な区分を指針とすること

表7 アクター・ファクター・マトリクス(大気汚染対策)

役割	ファクター			
	政策・対策遂行(知識実践)	政策・対策評価(知識再生産) 政策・対策形成(知識利用)	政策・対策課題設定(知識創造・蓄積)	
能力要素	P: 政策・対策の遂行能力	R: 環境対策資源の運用能力	K: 知識・技術の提供能力	
アクター	G: 政府	規制的手法 ・大気汚染対策に係る法規制の制定・運用 ・環境基準、排出基準の設定・監視 ・大気汚染対策条例、基本計画の策定・運用 ・立入検査等による監視 経済的手法 ・環境税、課徴金・補助金制度の整備・運用 自主的手法 ・公害防止協定(法的根拠なし)の締結 ・環境学習・教育の推進	予算 ・大気汚染対策予算の拡充 人員・組織 ・環境関連部署の設置 ・大気汚染対策組織(委員会、審議会、企業・市民協議会)の設置 ・大気汚染対策部署職員数の拡充 施設・設備等 ・大気汚染モニタリングシステムの整備 ・大気汚染警報設備の設置、情報システムの整備	調査・研究 ・大気汚染の要因・メカニズム等の研究 ・大気汚染対策技術の開発、ノウハウの蓄積 ・大気汚染対策の政策研究 情報公開・共有 ・大気質実態、大気対策情報の公開 ・職員教育・研修の実施
	F: 企業	規制的手法 ・大気汚染対策に係る法規制の遵守 ・環境基準、排出基準の遵守 ・大気汚染対策条例、基本計画への対応 経済的手法 ・補助金制度等の活用による対策 自主的手法 ・公害防止協定(法的根拠なし)の締結 ・財・サービス生産等、全プロセスでの環境負荷抑制 ・ISO14001の取得、ESCO事業の導入等	予算 ・大気汚染対策予算の拡充 人員・組織 ・環境関連部署の設置 ・環境関連部署職員数の拡充 ・環境管理者数、公害防止管理者数の拡充 施設・設備等 ・エンド・オブ・パイプ設備の導入 ・クリーナー・プロダクション技術・設備の導入	調査・研究 ・工場・事業所の自社モニタリング ・大気汚染対策技術の開発、ノウハウの蓄積 情報公開・共有 ・環境報告書、環境会計の作成・公開 ・職員教育・研修の実施
	C: 市民社会	規制的手法 ・大気汚染対策に係る法規制の遵守(野焼き等) 経済的手法 ・補助金制度等の活用による対策(NGO、NPO) 自主的手法 ・苦情、要望、ロビイング ・エコドライブ等、省エネ・省資源生活への転換 ・グリーン購入、エコファンド	予算 ・環境対策予算の拡充(NGO、NPO) 人員・組織 ・環境対策人員数の拡充(NGO、NPO) ・NGO、NPO活動への参加 施設・設備等 ・環境対策に係る施設・設備の確保(NGO、NPO) ・環境配慮製品の活用(省エネ、新エネ機器)	調査・研究 ・調査・研究(NGO、NPO) ・監視・モニタリング(NGO、NPO) 情報公開・共有 ・大気汚染状況の把握 ・環境学習、環境教育の実施

(出所)筆者作成

で、能力要素に係るデータの体系的な収集・選択が容易となる。

2点目は、効率的かつ効果的な評価実施である。UNDP(1998)では、評価フレームワークの曖昧さ、あるいは柔軟さにより、個々の環境問題ごとに、膨大なチェックリストを作成することになる。しかも、そのチェックすべきとして挙げられた能力要素は、能力要素の関係性や政策サイクルにおける役割が不明であり、他の評価基準との関連性が見えにくい、独立的もしくは重複性の高い評価基準に陥りやすい。このことは、全てのチェック項目が政策成果に貢献する適切な評価基準であるか、また、チェックの必要性があるかは確認できない。アクター・ファクター分析は、政策成果への貢献が確認された、評価すべき能力要素を評価

基準としており、評価に係る労力と効果の観点で優位である。このことは、途上国のオーナーシップ、および自己診断能力の形成に向けた、途上国自らの評価実施を可能とする簡便な評価方法としても適当である。表7では、大気汚染対策のマトリクスを例示しているが、政策サイクル、知識サイクル、能力要素という、一般的な基準で能力を区分し、下位の評価項目も標準的な項目区分としているため、他環境分野にも応用できる共通的な評価フレームワークになり得る。

また、実証分析の表3、4で、能力要素ごとのウェイトを示したが、政策課題設定、政策形成・政策評価、政策遂行の3つの役割の両市の各ウェイトは、4.2と10.7%、42.6と52.0%、28.3と35.2%であり、2都市の平均は、7.5%、47.3%、31.8%とな

る。大気汚染政策に係る能力の都市間比較、また、時系列比較を行う目的であれば、能力要素のウェイト付けによる定量的な評価も可能である。このウェイトは、日本での都市構造および大気汚染対策の異なる都市でも大きな差は無く、ある程度、一般的な数値として解釈可能である。

3点目は、能力形成に向けた適切な政策提案である。UNDP(1998)での個人、組織、社会システムの関係性は、包括性のある階層関係になっているため、階層ごとの能力評価の結果からは、どこに資源投入するのが望ましいかは判断できない。アクター・ファクター分析のフレームワークは、政府・企業・市民というアクター区分と、能力要素(政策サイクル、知識サイクル)で区分される重なりが無い9つのセルにより、誰が、何をすべきかの役割分担が明確になっている。したがって、能力評価結果に基づき、アクター間の代替可能性を踏まえ、どのセルに資源投入すべきか明確であり、能力形成に向けた、より実効性の高い政策提案が可能となる。

ただし、社会的環境管理能力は、社会経済状態、環境質との相互関係の中で位置づけられるものであり、実際の評価に際しては、社会的環境管理能力のみに着目するのではなく、これら相互関係に基づくトータルシステムとしての視点も合わせて必要となる。

5. 結論

本研究は、行政の大気汚染政策に係る能力要素を体系的に抽出し、能力要素の大気質改善への貢献とウェイトを実証的に確認することで、能力の評価基準を示した。そして、その評価基準を、政策サイクルと関連付けた能力評価に係るフレームワークの核に位置づけ、新たな能力評価手法として、アクター・ファクター分析を提案した。本研究により、以下のことが明らかとなった。今後は、モデルケースでの評価実施などを通じて、実践的な評価手法の確立に向けて、更なる知見を蓄積していく必要がある。

(1) 大気汚染政策に係る能力の構成要素は、「科

学的知見」の提供能力、「環境政策資源」の運用能力、「直接規制」の遂行能力、「経済的支援」の遂行能力の4つで示される。

- (2) 4つの能力要素全てがSO₂濃度改善に寄与し、大気質の改善という政策成果に貢献する能力要素であることが確認できた。また、ウェイト付けされた能力要素により統合化された大気汚染政策に係る能力が、SO₂濃度改善に貢献すること、また、日本の環境政策の発展段階区分と整合的であることから、4つの能力要素のウェイトの妥当性も確認できた。これにより、4つの能力要素が能力の評価基準として適切であることが示された。
- (3) 政策サイクルにおいて4つの能力要素が担う役割を設定することで、評価の視点の多様化と能力概念の具現化を可能とした能力評価に係るフレームワークを設定した。
- (4) 新たな能力評価手法として、アクター・ファクター分析および分析のフレームワークとしてのアクター・ファクター・マトリクスを示し、UNDPで示されている能力評価手法との比較を行い、評価基準の体系的な整理、効率的かつ効果的な評価実施、能力形成に向けた適切な政策提案の3点において優位性を指摘した。

付記

本論文は、21世紀COEプログラム「社会的環境管理能力の形成と国際協力拠点」(広島大学大学院国際協力研究科、拠点リーダー:松岡俊二)の研究成果の一部である。

注記

- 1 第37章「開発途上国における能力開発のための国のメカニズム及び国際協力」では、持続可能な発展は人々や制度の能力によって多くが規定される、と示されている。
- 2 Capacity in Environmentの定義は、「社会が環境問題を把握し、対処する能力(ability)」(OECD 1994)、「個人、集団、機関、制度が、与えられた状況において持続可能な発展を成し遂げるために必要な努力の一部として環境問題に取り組む能力(ability)」

- (OECD 1995)とされている。
- 3 政令指定都市は、都道府県からの一定の権限委譲を受けており、ある程度独自に政策立案・実施が可能であり、意思決定に際してのバイアスが少ない。大阪市では、1965年に公害防止条例が大幅に改正され、同条例施行規則の一部改正となった1967年に、条例権限の大部分が大阪市長に委任された。北九州市では、1970年に大気汚染防止法に係る県知事権限が委譲された。そして、北九州市大気汚染防止連絡協議会を設立し、対策が進められた。
 - 4 大都市統計協議会「大都市比較統計年表」(各年版)に拠る。
 - 5 藤倉(1998)は、これに関して、「(北九州)市は大気汚染防止連絡会議に所属する30社32工場を説得すれば、対策にかなりの効果が期待できた。その点は、西淀川区公害対策特別機動隊が域内一千余の中小・零細工場を相手にして、しらみつぶしに「どぶ板作戦」を行わざるを得なかった大阪市とは状況が異なる」としている。
 - 6 長島・新堂(2001)は同様の方法を用いて、「厚生指標」を作成している。
 - 7 大気環境学会史料整理研究委員会(2000)は、地方公害試験研究機関の大気汚染研究推進に果たした役割は大きいと評価している。また、日本の大気汚染経験検討委員会(1997)は、公害対策に係る市立の研究所のみを対象としてはいないが、「企業の取組が十分でなかった公害問題の初期においては、測定技術、対策技術の開発等にも地方公共団体の技術者が大きな役割を果たした」と評価している。
 - 8 衛生費は、環境対策費に加えて、保健、健康に係る費用も含むが、1970~2000年の全期間を通じて、環境対策のみの予算を表すデータはないため、これに代替した。
 - 9 能力(Capacity)は、フローではなくストックで示されるべきという見方もある。松岡・本田(2002)は、Sen(1985)、UNDPの人間開発論を踏まえ、Capacityをパフォーマンス・ベースの能力、Capabilityを人間活動の選択の自由・拡大につながる潜在的な能力としている。そして、その関係性を、「所与の条件での実践能力であり、その成果」であるCapacityと、その根底にある潜在能力としてのCapabilityと位置づけている。本研究での大気汚染政策に係る能力は、実際の活動として現れたパフォー

マンスとしての能力(Capacity)を計測しており、年度ごとのパフォーマンスに対するフローにおいて能力を捉えている。もちろん、フローとしての能力(Capacity)は、その時間展開に基づきストックとして形成された潜在能力(Capability)の水準に基づき発揮されるものである。

- 10 大気環境学会史料整理研究委員会(2000)は、大気汚染研究は、政治、行政に影響を与え、汚染低減の達成に寄与したかによってのみ評価されるべきとしている。
- 11 植田(1996)にあるように、環境政策手段としての直接的な手段、間接的手段、基盤的手段という区分において、「経済的支援」は間接的手段の範疇に入る。ただし、あくまでも経済的優遇措置であり、環境税などの市場メカニズムを用いた政策手段とは性質は異なる。
- 12 北九州市は1968年に北九州市公害防止資金融資制度要綱を制定し、大阪市は1967年に公害防止設備資金融資要綱を設けた。
- 13 世界銀行では、融資業務を相手国のキャパシティビルディングの機能を果たすものとして捉えている(ヒアリング)。この場合、行政の能力とともに、企業の能力形成にも貢献する政策制度であるといえる。

参考文献

- 植田和弘(1996)『環境経済学』、岩波書店
 上山信一(1998)『「行政評価」の時代』、NTT出版
 大阪市(1994)『大阪市公害対策史』、地球環境センター
 勝原健(2001)『東アジアの開発と環境問題』、勁草書房
 北九州市(1998)『北九州市公害対策史』、『北九州市公害対策史 解析編』、北九州市
 JICA(2004)『キャパシティ・ディベロップメント・ハンドブック：JICA事業の有効性と持続性を高めるために』、JICA
 大気環境学会史料整理研究委員会(2000)『日本の大気汚染の歴史』、LATTICE
 寺西俊一(1994)「日本の環境政策に関する若干の省察：アジアNIEsへの教訓として」、小島麗逸、藤崎成昭編『開発と環境：アジア「新成長圏」の課題』、アジア経済研究所、203-227
 長島直樹・新堂精士(2001)「新たな「厚生指標」作

- 成の試み」、『ESP』、426 : 68 71
- 日本の大気汚染経験検討委員会編 (1997) 『日本の大気汚染経験 : 持続可能な開発への挑戦』、ジャパンタイムズ
- 原嶋洋平・森田恒幸 (1995) 「東アジア諸国の環境政策の発展過程の比較分析」、『計画行政』、18 (3) : 73 85
- 藤倉良 (1998) 「公害対策の社会経済的要因」、北九州市 『北九州市公害対策史 解析編』、北九州市、185 247
- 藤倉良 (2002) 「日本の地方公共団体の硫黄酸化物対策」、寺尾忠能・大塚健司編 『「開発と環境」の政策過程とダイナミズム』、アジア経済研究所、37 78
- 松岡俊二・松本礼史・河内幾帆・岩瀬誠 (2000) 「東アジアにおける社会的環境管理能力の形成について」、『東アジアへの視点』、2000年秋季特別号 : 92 105
- 松岡俊二・本田直子 (2002) 「環境援助における能力開発とは何か : 環境管理能力の形成 (CDE) 概念のレビュー」、『国際開発研究』、11 (2) : 149 172
- 松岡俊二・朽木昭文編 (2003) 『アジアにおける社会的環境管理能力の形成 : ヨハネスブルグ・サミット後の日本の環境ODA政策』、アジア経済研究所
- 松岡俊二・岡田紗夏・木戸謙介・本田直子 (2004) 「社会的環境管理能力の形成と制度変化」、『国際開発研究』、13 (2) : 31 50
- 真山達志 (2001) 『政策形成の本質』、成文堂
- 山内康英・鈴木寛・渋谷修一 (2000) 『政策決定の新しいデザインと「知識マネジメント」』、『GLOCOMreview』、5 (5) : 1-17
- Barro, R. J. (1991). Economic Growth in a Cross Section of Countries, *Quarterly Journal of Economics*, 106, 407 443.
- Boesen, Jannik, and Lafontaine, Alain (1998). *The planning and monitoring of capacity development in environment (CDE) initiatives*, CIDA.
- Honadle, B. W. (2001). Theoretical and Practical Issues of Local Government Capacity in an Era of Devolution, *Journal of Regional Analysis and Policy*, 31 (1), 77 90.
- Janicke, Martin, and Weidner, Helmut (eds) (1997). *National environmental policies : a comparative study of capacity-building*, Berlin: Springer.
- Lavergne, R al, and John, Saxby (2001). Capacity Development: Vision and Implications, CIDA Policy Branch, *Capacity Development Occasional Paper Series No. 3.*, CIDA.
- Mulligan, C. B., and Sala-i-Martin, X. (1997). A labor income-based measure of the value of human capital: An application to the states of the United States, *Japan and the World Economy*, 9 (2), 159 191.
- OECD (1994). *Capacity Development in Environment : proceedings of a workshop held in Costa Rica. 9 11 November 1993*, Paris :OECD.
- OECD (1995). *Developing Environmental Capacity : a framework for donor involvement*, Paris :OECD.
- OECD/DAC (1999). Donor Support for Institutional Capacity Development in Environment: Lessons Learned, *Evaluation and Effectiveness No. 3.*
- Sen, Amartya (1985). *Commodities and capabilities*, Amsterdam : Elsevier Science.
- UNDP (1998). *Capacity Assessment and Development in a Systems and Strategic Management Context*, UNDP/BDP/Management Development and Governance Division
- UNDP/GEF (2003). Capacity Development Indicators, *UNDP/GEF Resource Kit (No. 4)*
- UNEP/WHO (1996). *Air Quality Management and Assessment Capabilities in 20 Major Cities*, London : MARC.

(2005.12.28受理)

Evaluation of Social Capacity for Urban Air Quality Management

Kazuma Murakami

Shunji Matsuoka

Graduate School for International Development and Cooperation,
Hiroshima University
mkazuma@hiroshima-u.ac.jp smatsu@hiroshima-u.ac.jp

Abstract

Since early 1990s, Capacity Development and Capacity Development in Environment (CDE) are discussed in the field of aid assistance. In the current literature, although there is consensus on necessity of capacity development, there is none on its definition as well as its measurement. Regarding this, this study tries to develop a new evaluation method for capacity development. As a case study, we examine the elements of governmental capacity for urban air quality management in major cities of Japan. In this respect, city-level panel data in 1970-2000 period are used in the analysis. By using Factor analysis, we estimate four elements of capacity that contribute to urban air quality improvement. These four elements are further arranged into three factors. Using these factors, we propose an evaluation form called as Actors-Factors matrix. Evaluation method based on Actors-Factors matrix is superior to UNDP's method in a systematic classifies of evaluation criteria, efficient and effectives evaluation, and appropriate policy design for capacity development.

Keywords

capacity development, evaluation, indicator, factor analysis, air quality management

【研究論文】

インドネシア前期中等教育における地方分権化の影響 教育環境と教育成果に関する因果モデルによる評価

戸井 敦子 牟田 博光

東京工業大学大学院

atoi@hd.hum.titech.ac.jp

muta@hum.titech.ac.jp

要 約

インドネシアでは、2001年の地方自治法の施行を受け、中央から県・市への予算直接配分や地域参加型の学校運営の導入など、教育分野における地方分権化が進められている。

本稿では、インドネシアの前期中等教育に着目し、地方分権化実施の前後において、教育環境が教育成果に与える有効性がどのように変化したかを、因果モデルを用いて検証した。

主な分析結果として、学校の予算状況の改善は、教員資質を高めるなど教育環境の向上をもたらし、学力の向上や中退率の減少という教育成果の改善に寄与するが、地方分権化後はこれらの因果関係について影響力と有効性に低下が認められる； 予算規模の大きい学校グループと小さい学校グループを比較すると、地方分権化によるマイナスの影響は予算規模の小さい学校グループにより多く現われていることが明らかになった。

地方分権化が教育現場に与える影響について、更に継続観察するとともに、教職員の質向上や財政的措置により、地方分権化後に拡大が認められる地域や学校間の格差の是正に取り組んでいくことが重要である。

キーワード

地方分権化、インドネシア、前期中等教育

1. はじめに

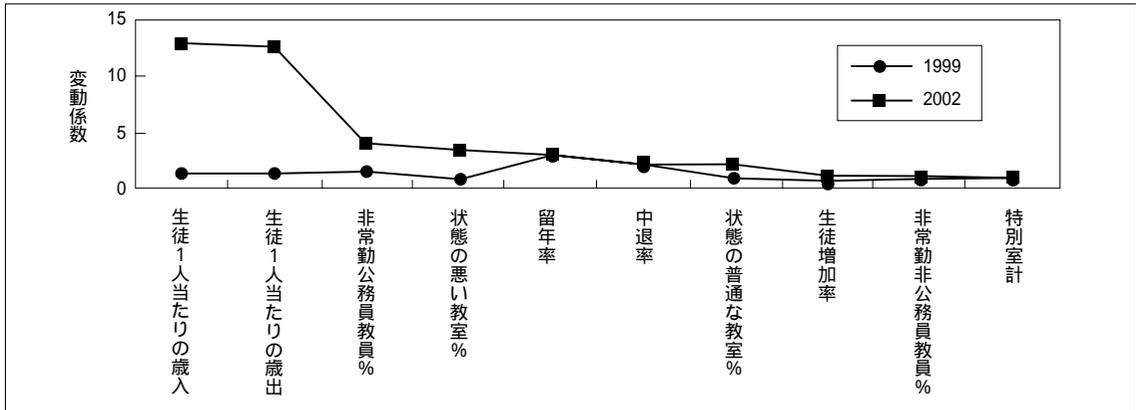
地方分権に関わるインドネシアの概況

2001年の地方自治法の施行により、インドネシアにおいて地方分権化は本格的に実施され始めた。分権化の背景には、旧スハルト時代の負の遺産である汚職や業務遂行の非効率性などの課題を抱える中央集権体制から、新しい時代に則した行政システムとしての分権化を完成し、更なる発展を目指すというインドネシアのねらいがある。こ

のような分権化への移行は時代の潮流でもあり日本もその例外ではない。

しかし、地方分権化が常に問題解決の万能策にはなりえないことは、様々な先行研究が明らかにしているところである。1980年代に新自由主義政策の一環としてチリ政府が実施した教育行政の地方分権化は、教員の非公務員化や市場原理の導入で、効率性の向上を図ることをベースにしていたが、中等教育の就学率増加など量的拡大をもたらした一方で、学校類型間での階層間格差や学力の

図1 変動係数の年度別比較



(出所) 学校基本調査結果から筆者作成

低迷などの副作用ももたらした(斎藤2004)。ブラジルでは1994年に始まったカルド・ゾ政権の教育改革によって初めて地方分権化を導入し、市や学校への教育運営責任の移譲や財政改革によって、学校の自治力・教育力の強化を図り、成功事例も報告されている一方で、北東部など貧しい地域では改革の遅れが指摘されている(江原2005)。

インドネシアの地方分権化は、「地方自治体が、公共サービスの提供及び政府開発事業の有効性と効率性を向上させ、地方自治体と住民の自立性を高め、民主主義を育成する」ことを目標とし、主に県・市の自治権拡大を図っている。教育改革に関しては、教育の質、平等性、妥当性、効率性の向上を方針に掲げ、それまでの教育行政組織の大幅な改革に臨んでいる。代表的な組織改編としては、初等及び前期中等教育は県・市の教育局に、後期中等教育は州政府の教育局に教育行政上の権限が委譲されるなど、国家教育省の役割と権限の大幅縮小、及び地方政府や組織の権限の拡大が行われた。また、教育行政の一元化として、それまで地方に置かれていた国家教育省の州、県、郡事務所が廃止された。

加えて、学校に基礎をおいた経営(School-Based Management、以下SBM)と地域社会に基礎をおいた教育(Community-Based Education、以下CBE)が、初等中等教育の運営システムを改善するための中心的施策として推進されている。

広範囲な群島からなるインドネシアは、地域間

の格差などの課題を抱えながらも、強力な中央集権体制の下で統合されてきた。新しいシステムが従来からの課題を克服しつつ発展していくことが望まれるが、新たな課題が生じることも懸念されている。JICA(2005)の報告書でも、分権化が実施された2001年度の調査結果として、中部ジャワ州や北スラウェシ州での就学率の低下や、県や都市間での就学率格差の拡大が指摘されている。

図1は国家教育省管轄の公立中学校全体の教育データについて、分権化後(2002/2003年度)における変動係数値が最も高かった上位10項目(以下、変数)を取り上げ、分権化前(1999/2000年度)と比較したものである。

変動係数とは、標準偏差と平均値の比であり、歳入や中退率など測定単位の異なる変数や、測定単位が同じでも分権化の前後で中心が著しく異なる変数においても、データの散らばり具合を相対的に比較することができる。分権化前と比較すると分権化後は全体的に変動係数が高くなっており、データ間つまり学校間の散らばりが大きくなっていることがわかる。この傾向は図中の10変数以外の教育環境においても認められた。公立中学校は他の学校種よりも比較的教育環境が整っているが(戸井・牟田2003)、それでも分権化後の教育環境については、データ間つまり学校間の差異が広がっていること、特に予算に関する変数(生徒1人あたりの歳入・歳出)において学校間格差が著しくなったことがわかる。

この現象の背景にあるものとして、財政分権化による地方の財政状況の変化があげられるが、それに関しては松井(2003)が詳しく論及している。インドネシアの地方財政は、植民地時代より中央からの移転資金(歳入分与、一般配分金、特別配分金から成る)に大きく依存してきた。この依存体制は、各自治体の自発的な発展の妨げになっていたが、地域間格差の大きいインドネシアにおいて、格差是正の重要な役割も果たしていた。しかし、分権化後には、歳入分与における国民1人当たりの配分格差は2倍に広がっている。

この歳入分与の格差拡大の背景のひとつには、分権化を機に地方行財政の運営に変更が施されたことがある。例えば、中央集権時代にその富みのほとんどを中央に徴収されてきた東カリマンタン等の資源富裕地域においては、分権化以前に比べると優遇的な配分が行われるようになった。歳入分与による地域間格差の拡大を緩和させる役目を有するのは一般配分金であるが、一般配分金の算出には前年度を下回らないための「均衡部分」が加味されており、結果的に格差是正の役割は十分に機能していない。地方分権化では地方の財政的自立が期待されているが、県や市レベルについていえば、それまで重要な税源であった地下水税や湧水利用税の州税化などにより、安定した税源がなくなったため、その自己財源比率は低下し、中央からの移転資金に依存せざるをえない構造となっている。移転資金の中身として、分権化以降は歳入分与の比重が高まっており、経済活動や資源賦存の偏りを考慮すれば、中央からの資金の配分差が短期的には地方間格差の助長をもたらすことになったと考えられる。また、分権化によって初等・前期中等教育の直接的所轄は国家教育省から県・市教育局に移管され、予算の交付先も県・市への直接交付となったため、交付される資金の用途や配分については県・市議会によって決定されることとなった。しかしながら、供給すべき地方公共財・サービスについての能力や関心が不十分なまま、以前よりも拘束のない移転資金が始まったため、地方政府の議会対策や備品購入など住民生活とは関連のない用途に使われる傾向も少なからず指摘されている。図1に認められたような予算の学校間格差の背景には、以上のような分権化

による地方行財政の変化や過渡期における混乱等が作用していると推察される。

インドネシアの前期中等教育

インドネシアにおいては、初等教育の就学機会が80年代に大幅に拡大し、完全就学がほぼ完成されている。しかし、中学校にはまだ課題が多く、粗就学率は70%台(百瀬2001)に増加してはいるが地域差が存在している。教育の質向上についても1999年に決定された国策大綱(1999~2004年)では基礎教育分野における最優先課題の1つとして重要視されているが、分権化にともなう変化が、中学校における教育活動にどのようなインパクトを与えたかを評価する論及は少ない。インドネシアにおけるSBMの一環である「学校を基盤とする質向上のための経営」(Manajemen Peningkatan Mutu Berbasis Sekolah:MPMBS)という施策に焦点を当て、学校内外における様々な問題を抱えながら、MPMBSが質向上に向けてどのようなしかけをしているのかについての研究(中矢2005)では、首都のある優秀校における成功事例が報告されているが、広範囲の中学校に与えたインパクトについては論及されていない。行政体制の変化が、それ以降の教育活動に影響した例としてペレストロイカ以降のロシアについての研究がある。森岡(2000)は、体制転換以降、学校の民主化、自治の拡大が叫ばれ相次いで改革案が出されたが、90年に行われた意識調査によって教育に対する国民の満足度はかなり低いこと、その理由として8割近くが物資・財政的基礎の弱体化を上げたことを報告している。改革に伴う実質的保障が十分になされないことによって、教育における経営・財政面の不均衡による地域及び学校間格差の拡大を懸念しているが、実際の教育成果へのインパクトは検証されていない。

教育成果の規定要因に関しては、生徒の家庭環境等の社会経済的格差による影響が大きいとする説(Coleman 1966;Bowls and Gintis 1976)や学校要因による影響を唱える説(Edmonds 1979; Psacharopoulos and Woodhall 1985)などがある。また、先進国に比べて教育に関わるリソースがより限られている途上国においては、学校環境の整備が教育成果の向上に大きな影響力をもつ事例も

報告されている(Cass & Deaton 1999; Fuller 1987)。

図1で認められたような教育環境における変動が、中学校における教育活動に及ぼした影響を明らかにすることは、分権化にともなう新たな取組みの必要性を探るためにも重要と考える。

そこで本研究では、インドネシアにおける地方分権化の影響を、教育環境と教育成果の因果関係の変化に着目し考察していく。中学校における主要な教育環境要因を特定した上で、要因間の関連を検討し、教育環境のどの要因が教育成果の向上に影響力を持つのかを検討する。そして、それら教育環境と教育成果の因果関係が分権化の前後でどのように変化したかを検証していきたい。具体的には、1)現象の因果関係を調べるための統計的手法である共分散構造分析によりモデルを構築し、インドネシアの中学校における教育環境と教育成果について、要因間の因果関係を検討する。その上で、それらの関係が分権化によってどのような影響を受けたのか全体的傾向を把握する。次に、検証すべき現象を踏まえた上で、2)分権化後に著しい学校予算の格差拡大に着目し、予算の多いグループと少ないグループに大別して分析を行い、分権化の影響は学校の予算状況によって異なるかを検討する。

インドネシアのフォ・マルな中学校には、国家教育省管轄のSLTP (Sekolah Lanjutan Tingkat Pertama) (公立、私立)と宗教省管轄のMTs (Madrasah Tsanawiyah) (公立、私立)がある。それぞれのタイプの学校では、学校予算の財源、教員人事など学校組織や運営方法は異なるため、本研究では、このうち学校数ではインドネシア全中学校の34%、生徒数においては58%と最大の割合をしめる (Department Pendidikan National 2003) SLTP公立校を分析の対象とする。

2. 分析1: SLTP公立校全体にみる地方分権化の影響

(1) 目的

SLTP公立校について、教育環境と教育成果との関係を示す因果モデルを構成し、その妥当性の

検討を通して、要因間の因果関係を探る。そしてインドネシアの地方分権化がそれらの因果関係にどのような影響を及ぼしたかについて、全体的傾向を把握する。

(2) 方法

分析に使用するデータ インドネシアの国家教育省が管轄する全中学校に対して毎年実施する学校基本調査データの原表を用いる。地方分権化以前の状態を示すデータとして学校基本調査の1999/2000年度版を、分権化後のデータとして同調査の2002/2003年度版を使用した。欠損値のあるケースを除外した1999/2000年度データのSLTP公立校は4853校、2002/2003年度データでは4045校であり、全SLTP公立校のそれぞれ46.2%、36.9%であった。

因果関係モデルの構成 まず、中学校における主要な環境要因を探るため、学校基本調査に含まれる中学校関連の全項目 (観測変数) を投入して探索的因子分析 (主因子法、斜交プロマックス回転) を行い、これら観測変数の背後に潜んで特に深い影響を与える主要な因子 (潜在変数) を抽出した。教育要因として5つの因子が抽出され、因子負荷量の大きさ及び観測変数の内容に基づいて、それぞれ「校長資質」、「教員資質」、「学校予算」、「学習環境」、「教育成果」と命名した。

教育環境と教育成果の関係については、知識や経験など教員の質的要因の学力への影響 (Fuller, 1987, Jimenes & Paquo, 1996) や、生徒対教員の人数や設備環境への投資による教育効果への影響 (Cass & Deaton, 1999) などに関する先行研究がある。

インドネシアでは、校長は教育活動、学校経営、教職員の育成 (研修・人事異動の申請等)、土地・施設・設備の管理補修についての責任を持つと規定されており (中矢2001)、SBMやCBEの導入によって、分権化後は校長の役割が一層重要性を増すと推測される。中学校の教員資格は最低3年間の高等教育が必要とされているが、その資格を充たしているのは約半数にすぎず質の向上が課題となっている (百瀬2001)。また、教員不足も

地域によっては深刻であるが、予算要求に関して中央集権時代は中央の予算配分次第であったので、学校の早急のニーズは反映されにくく、要求がすべて満たされる訳ではなかった。教員不足の問題を抱える学校は保護者からの寄付金によって安く働いてくれる非常勤教員を探さなくてはならず、学校間に格差が生じていた。特に、新設校や遠隔地は、教員配属の遅れや定着率の悪さから問題は更に深刻であった(中矢・諏訪2001)。図書館や理科実験室のない学校は3割近く存在するとされており、施設・設備の学校間格差も課題とされている。インドネシアの教育予算が国家予算に占める割合は分権化前でも7%台とタイなど近隣諸国と比べても低く、慢性的な予算不足状態にあったものが、分権化後は5%台に落ちている(Ministry of National Education,2003)。

これらの事前分析や先行研究、そしてインドネシアの教育事情に基づき、抽出された5因子の最初の2因子は学校環境のうちの人的環境を、続く2因子は物的環境を、最後の因子は教育成果を示し、因果モデルを構成する概念(構成概念)に適すると判断し、次の2つの仮説からなる因果モデルを構築してSLTP公立校における教育環境と教育成果との因果関係と、地方分権化後に生じた変化について検証することにした。

1) SLTP公立校において、生徒の学力や進級といった教育成果は、主要な教育要因として抽出された4因子によって直接・間接的に規定されている。予算が潤沢な学校であれば、質の高い校長を得やすく、更に教員集団の質も高まっていく。また、予算が潤沢であれば教材・設備などの学習環境も整備されていく。生徒を指導する教員集団の質が高く、学習環境が整備された学校ほど生徒の学力も高く、中退率も低くなる。

2) しかし、分権化という新しい体制への移行期にある学校現場では、変化に伴う混乱や学校間の予算格差の拡大などが影響し、教育環境から教育成果への影響力や有効性が分権化前と比較して減じたのではないかと。そのため新たな取組みの必要性が生じているのではないかと。

以上の仮説をパス図によって表したものが図2である。煩雑さを避けるために、観測変数や潜在変数(構成概念)に関連した誤差変数の表示は省

略した。因果モデル内の各概念を構成する観測変数は表1の通りである。なお表1にOutputsとして記載されている卒業試験と中退率の2つの観測変数は、元々は事前の因子分析で抽出された「教育成果」因子を構成する観測変数であったが、この因果モデルでは、卒業試験結果を生徒の学力を示す指標、中退率は進級状態を示す指標ととらえ、別々のアウトプットとして投入することによって、教育成果をより具体的に検証することとした。

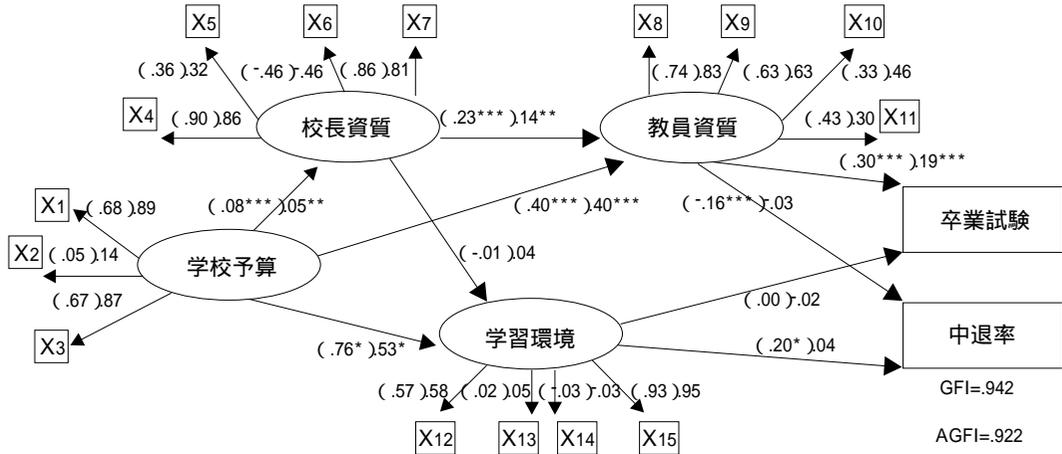
分析方法 SLTP公立校の教育環境と教育成果について想定した因果モデルの妥当性と、その因果関係が地方分権化によって受けた影響を共分散構造分析(豊田 1998; 山本・小野寺 1999; 田部井 2001)により検討する。2つの異なる年度を比較するため、分析にあたっては各概念変数に等値制約を施した多母集団の同時分析を用いた。多母集団とはこの場合1999/2000年度と2002/2003年度の2つの年度集団のことである。この2つの母集団において共通の潜在変数(因子)や因果関係が存在すると仮定する場合、個別に分析し比較するのではなく、多母集団の同時分析を行うことによってモデルの相等性や因子の普遍性などを検証することが可能になる。分析に際しては、異なる年度における因子の不変性を検証するため、各因子から観測変数へひくパスに等値制約を課している。分析にはAmos5.0を使用した。

(3) 結果

まず、因果モデルの全体的評価を行うため、GFI(Goodness of Fit Index:適合度指標)とAGFI(Adjusted Goodness of Fit Index:修正適合度指標)を求めた。GFIはモデルの説明力を表す指標であり、想定した因果モデルがデータをどの程度説明しているかを示す。AGFIは同様に説明力をあらわす指標であるが、推定するパラメータ数の影響を修正して算出しているため、より安定して説明力を量ることができるとされている(田部井 2001)。結果は図2の右下にある通り、GFIは.942、AGFIは.922といずれも高い値を示し、モデルとデータの適合度は高く、想定した因果モデルはデータをよく説明していると判断された。

図2 SLTP公立校における教育環境と教育成果の因果モデル

(カッコ内 = 分権化前、カッコ外 = 分権化後、数値は標準化解)



(出所) 筆者作成

***p<.001 **p<.01 *p<.05

表1 分析に用いられた潜在変数と観測変数

潜在変数	観測変数	観測変数の定義
Outputs	卒業試験 (10段階評定)	中学校3学年が受ける国家卒業試験の各学校平均点
	中退率 (%)	在籍生徒数に対する中途退学者の割合
学校予算	X1. 政府資金	政府から交付される学校運営費 (生徒1人当たりRp.)
	X2. 繰越金額	前年度繰越金等の収入 (同上)
	X3. 教員給与資金	常勤教員の給与資金 (同上)
校長資質	X4. 勤務年数	校長の勤務年数
	X5. 等級	校長の等級 (1~18級)
	X6. 学歴	校長の学歴 (0~17段階)
	X7. 年齢	校長の年齢
教員資質	X8. 平均勤務年数	各学校の教員の平均勤務年数
	X9. 常勤率	全教員に占める常勤教員の割合
	X10. D1率	学歴がディプロマ1 (1年間の高等教育) 教員が占める割合
	X11. 4級教員率	等級が4 (最高級) の教員の割合
学習環境	X12. 特別教室	図書室、実験室など特別教室の合計数 (生徒1人当たり)
	X13. 教材・備品	教科書、教材、備品への年間支出 (生徒1人当たりRp.)
	X14. 保護者寄付	施設の修理・補修、非常勤手当等に使用される保護者寄付金 (同上)
	X15. 生徒対教員	生徒1人あたりの教員数

(出所) 筆者作成

この因果モデルの部分的評価については、1) 因果モデル内の構成概念から観測変数へひかれる影響指標はすべて5%以下の水準で有意であり、2) 影響指標の中には.02など低いものもあるが、大部分は.30より高い数値を示している、の2点から、構成概念と観測変数の関係は適切に対応していることが示された。

次に、求められた構成概念間の因果係数から、モデルにおける要因間の因果関係と、その因果関係が地方分権化によってどのような影響を受けたかを検証した。構成概念間の矢印上に記載されている2つの因果係数のうち、カッコ内は地方分権化政策が実施される前の数値、続くカッコ外の数値は分権化後のものである。因果係数に付随する星印はその関係が統計的に有意であることを示している。星の数が多いほどその有意性は高く(0.1%以下の水準で有意な場合は***、1%以下の水準では**、5%以下の水準では*)、星印がついていない因果係数は統計的に有意でないことを表す。

まず、分権化前の状態を見てみると、教育成果を向上させる直接要因として、人的要因である「教員資質」が重要な役割を担っていることがわかる。「教員資質」から「卒業試験」及び「中退率」への因果係数は、それぞれ.30と-.16であり、有意性も高い。表1に示した「教員資質」に関する観測変数の内容と、モデル内の各影響指標の値を考え合わせると、経験豊かで等級も高い教員が常勤としてたくさんいるほど生徒の学力が向上し中退も減少することが分析から示唆された。一方、物的要因である「学習環境」からは「中退率」にのみ因果関係が認められるがその有意性は高くはない。

「教員資質」の向上に影響するのは「校長資質」と「学校予算」である。両方とも有意性は同じように高いが、因果係数値では「校長資質」の.23に対して、「学校予算」は.40とより高く、学校予算の潤沢な学校ほど「教員資質」も向上することが示された。また、校長については、学歴の高さよりも、経験豊富で等級も高い校長の存在が教員集団の資質向上に影響することが示された。また、「校長資質」は「学校予算」から影響をうけるが、その因果係数は.08と比較的低く、「学校予算」が

「校長資質」を規定する力はあまり大きくないことが明らかになった。

「学習環境」については、「学校予算」からの因果係数が.76と高いが、その有意性は低い。

次に、分析で明らかになった以上のような因果関係が、分権化後にどのように変化したかを見ていきたい。「教員資質」は分権化後においても「卒業試験」の結果を向上させるのに有効な要因であるが、その因果係数は値を.30から.19へと減少させている。また、「中退率」への有意な影響力は認められなくなっている。

「教員資質」が「校長資質」から受ける因果係数も.23から.14と減じている一方、「学校予算」から受ける因果係数は変わっていない。「校長資質」が「学校予算」から受ける因果係数は.08から.05へと大きな変化はないが、有意性は低くなっている。

「学習環境」に関しては、「中退率」への有意な関係をなくしており、「学校予算」からの因果係数も.76から.53に減じている。

これらの現象をまとめると、教育成果の向上には人的要因の「教員資質」が有効であること；学校予算が潤沢なほど「教員資質」が向上するが、「校長資質」も重要な規定要因であること；「学校予算」の潤沢さは「学習環境」の改善に影響しうる、という主要な因果関係が認められるが、モデル内の少なからぬ因果係数値やその有意性の減少から、地方分権化後、教育環境から教育成果への因果関係において、その有効性と影響力に低下が見られると判断される。

以上、分析1においては教育成果の向上に有効な教育環境要因とその構造を明らかにし、その因果関係が分権化後にどのように変わったかについて、全体的傾向をとらえた。次の分析2では、予算別のグループに分けて検証を進めることにより、全体の分析では見出せなかった分権化の影響の詳細な部分についてさらに明らかにしていきたい。

3. 分析2：予算状況別にみる分権化の影響

(1) 目的

学校予算は、重要な教育環境要因であるが、そ

の学校間格差は地方分権化後に拡大している。分析2においては、SLTP公立校を学校予算の状況別に分類し、分析1で明らかになった分権化後の教育環境と教育成果との因果関係の低下が、学校予算の状況によってどのように異なるのかを検証する。

(2) 方法

分析1で使用したSLTP公立校に関する学校基本調査データから、生徒1人当たり学校予算の歳入平均値を算出し、平均値に満たない予算を有する学校グループと、平均値を超える予算を有する学校グループに分類する。分析1と同様の因果モデルと分析方法を用い、予算グループ別に地方分権化の前後における変化を比較し、その影響の差異を検証する。(以後、平均値に満たない予算を有する学校グループを「予算の少ないグループ」、平均値を超える予算を有する学校グループを「予算の多いグループ」と呼ぶ。)予算の少ないグループのSLTP公立校は1999/2000年度は3281校、2002/2003年度は2278校であり、予算の多いグループは1999/2000年度は1572校、2002/2003年度は1767校である。

学校が所属する地域の経済力を表すGRDP (Gross Regional Domestic Product、地域内総生産) に関して、2つのグループの平均値に統計的な有意差はなく、両グループとも、特定の経済レベルに偏ったサンプルではないことを事前に確認した。

(3) 結果

共分散構造分析による検討 予算状況別に分析したところ、予算の少ないグループについては $GFI=.942$ 、 $AGFI=.922$ 、予算の多いグループについては $GFI=.900$ 、 $AGFI=.865$ といずれも高い値が得られたことから、ともにモデルとデータとの適合度が高いことが確認された。前者の因果モデルは図3a、後者は図3bに示す。

まず、校長、教員といった人的要因については、予算の少ないグループでは「学校予算」が「校長資質」と「教員資質」を規定し、この関係は分権

化の前後ともかなり有意(0.1%以下の水準)である。一方、予算の多いグループでは、「学校予算」から人的要因への有意な影響力はあまり認められない。

「校長資質」から「教員資質」への影響は、分権化の前後を通して両グループとも有意であり、因果関係の値も安定している。加えて、「校長資質」から「学習環境」への因果係数が、分権化後に両グループとも有意となっている。

物的要因については、予算の多いグループでは、分権化の前後とも「学校予算」は「学習環境」と強く関係している一方、予算の少ないグループでは分権化前には有意であった「学校予算」から「学習環境」への因果係数は分権化後には有意性を失っている。

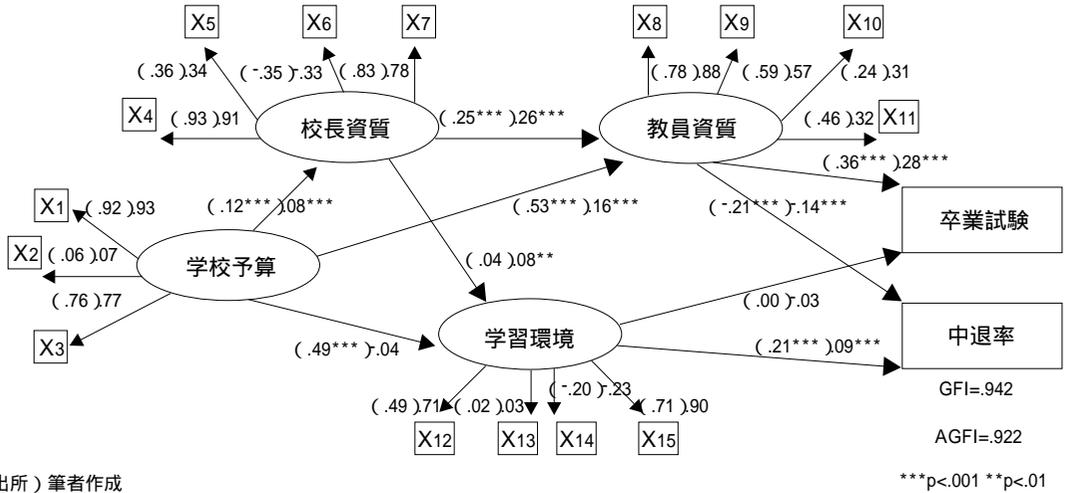
以上をまとめると、予算の多いグループでは、人的要因は「学校予算」によって規定されないが、物的要因は強く規定されており、この因果関係には分権化の前後を通して大きな変化はない。一方、予算の少ないグループでは、「学校予算」は人的要因を規定する重要な要因であるが、その影響力は分権化後に低下している。加えて、「学習環境」への有意性が分権化後には失われており、予算の少ないグループでは「学校予算」が他の教育環境に与える効果は減少している。しかし、「校長資質」の影響力については、分権化後も両グループにおいて安定していることに加えて、「学習環境」への有意性も現われている。

最後に、教育成果への効果については、SLTP公立校全体における分析結果と同様、「教員資質」の高さが、卒業試験結果の向上と中退率の減少の両方に有意に影響し、その効果は卒業試験結果に対してより強く現われている。「教員資質」が教育成果に及ぼす影響力は、分権化前はグループ間で大差はないが、分権化後は、予算の少ないグループの方でより顕著な影響力の低下が認められる。

「学習環境」については、両グループとも、「中退率」の増加への影響が示されている。ひとつの可能性は、学校規模の影響である。因果モデルを構築するための事前分析(因子分析)では、「学習環境」を構成する代表的な4つの観測変数の他に、生徒総数に関する変数も得られており、学校

図 3 a SLTP公立校における教育環境と教育成果の因果モデル

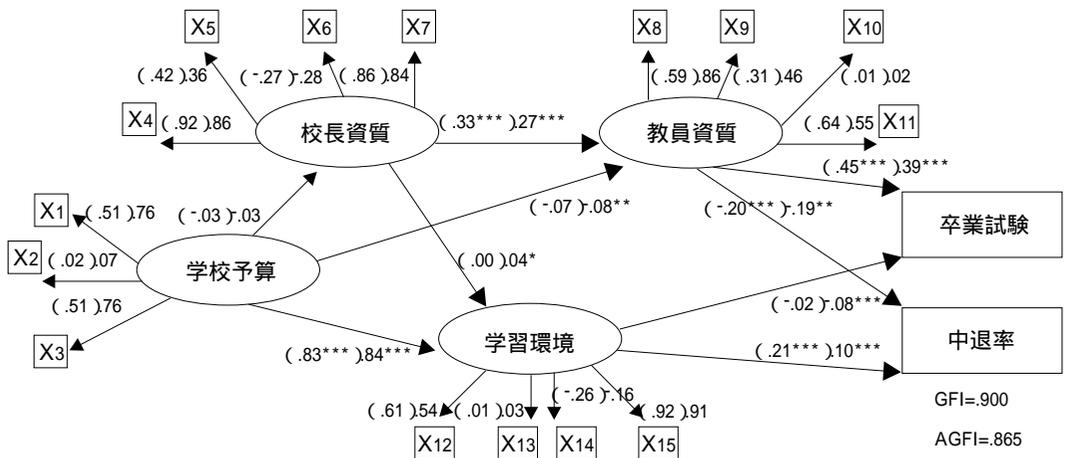
(予算平均に満たないグループ)



(出所) 筆者作成

図 3 b SLTP 公立校における教育環境と教育成果の因果モデル

(予算平均を超えるグループ)



(出所) 筆者作成

規模の小ささが「学習環境」を向上させることを示唆していた。しかし、小規模校の中には中退率の高い遠隔地の学校や都会でも生徒の集まらない不人気校も含まれる。「学習環境」の改善にもかかわらず「中退率」が増加となったのは、このような学校の存在が反映された可能性もある。い

れにしても、中退率の減少に「教員資質」以外の学校環境として何が有効であるかについては今後の課題である。

教育環境の水準比較 これまでは、要因間の因果係数の変化を基に分権化の影響を調べてきたが、

次はモデル内の構成概念に着目し、4つの教育環境要因（「学校予算」、「校長資質」、「教員資質」、「学習環境」）の状況が、分権化後どのように変わったか、そしてその変化には予算の多いグループと少ないグループでは差異が認められるのかを見ていきたい。田中・山崎・柳井・鈴木（1998）は、因果モデルについて、「同一モデルを異なる集団に対して適用し共分散構造分析を行った場合、因果係数による関係の比較を行うとともに構成概念の水準（平均）についても比較検討することが有用である。」と述べている。本稿でも田中らに従い、各構成概念に関して、予算の少ないグループと予算の多いグループ別に水準（平均）を求め、その差異を検討した。具体的には、構成概念は、因子を構成する観測変数の集合であるので、標準変換した各観測変数の合計点を概念得点とし、グループ別に平均値を算出し、その差異を各項目について比較した。結果は表2に示す。

分権化の前後を通じ、全体的に予算の多いグループの方が教育環境の水準は高い。分権化後の変化としては、予算の少ないグループでは「学校予算」及び「教員資質」において分権化の前後で統計的に有意な差異が認められる。予算の多いグループでは「学習環境」についてののみ有意な差異が認められる。この結果から、両グループとも「校長資質」に関しては、分権化による有意な変化は認められなかったが、予算の少ないグループでは「学校予算」と「教員資質」の水準が分権化後に低下し、予算の多いグループでは「学習環境」の水準が低下したと判断される。「教員資質」は予算の少ないグループにとって「卒業試験」と「中退率」の両方に影響する重要な要因であり、その

水準の低下は問題である。加えて、予算の少ないグループにおいては、「教員資質」を規定する「学校予算」の水準も低下している。分権化における教育環境水準の低下は、予算の少ない学校においてより深刻であると判断される。

4. まとめと考察

本研究では、学校現場における教育環境と教育成果の因果モデルを用いて、地方分権化がその因果関係に及ぼした影響を明らかにし、新たな取組みの必要性を探ることを目的とした。

分析1においては、SLTP公立校全体を対象とした分析によって、以下の2点が判明した。（1）学校の予算状況の改善は、校長や、特に教員集団の質向上に影響し、結果的に生徒の学力を向上させ中退も抑制することに貢献している。校長の資質も教員資質の向上に有効である。また、学校予算への投資は教材・施設・設備等の学習環境の整備にもつながる。（2）上記の主要な因果関係は分権化後も変わらないが、分権化の前後で比較すると、全体的に教育環境が教育成果に及ぼす影響力に低下が認められた。

そこで分析2においては、分権化後に拡大した予算の学校間格差に着目し、学校を予算の状況別に分類した上で因果モデルを検証し、分析1で得られた分権化の影響の更に詳細な部分を明らかにしていった。その結果、比較的予算状況の良いグループにおいては、分析1で得られたように、因果関係に低下が認められるものの、その減少の程度は若干であり、学校へのリソースの投入が生徒

表2 概念得点の平均値の比較

構成概念	予算の少ないグループ			予算の多いグループ		
	分権化前 (n=3281)	分権化後 (n=2278)	t値	分権化前 (n=1572)	分権化後 (n=1767)	t値
学校予算	- .81	-1.25	19.62***	1.70	1.62	.89
校長資質	.01	-.11	1.81	-.02	.14	-1.85
教員資質	-.58	-.81	3.22***	1.22	1.05	1.89
学習環境	-.70	-.71	.14	1.46	.91	5.48***

***p<.001

（出所）筆者作成

の成績や進級の改善に及ぼす効果は安定している。また、このグループの特徴としては、学校予算は学習環境を強く規定しているが、校長や教員など人的要因はあまり規定していない。校長資質の高さが教員集団の質を向上させ、それが生徒の学力や進級に効果を与えるという構図がある。校長資質は、分権化後、学習環境の向上にも有意な影響を及ぼすようになっており、重要な環境要因と判断される。

一方、生徒1人当たりの学校予算が平均に満たない比較的予算状況の悪いグループにおいても、分権化後、因果関係の影響に減少が認められたが、その程度はより顕著であった。このグループでは、予算を起点とする教育環境から教育成果への因果関係においては、一定の投入から生じる効果が分権化前より減少し、効用が小さくなっていると判断される。「校長資質」は、予算の多いグループと同様に重要な環境要因であることは、他の因果係数の値が低下したなかで校長資質から教員資質への因果係数は分権化後も安定しており、学習環境の改善にも有効性が認められるようになったことから確認できる。

最後に、各教育環境の水準の比較においては、(1)分権化の前後を通して、予算の多いグループの方の教育環境が良い、(2)予算の多いグループは分権化後、学習環境のみ水準の低下が見られるのに対し、予算の少ないグループでは学校予算と教員資質の2つに水準の低下が見られる、の2点が確認された。分析2において、予算の少ない学校では、学校予算は教員資質を規定する重要な要因であり、教員資質も卒業試験と中退率を改善するための重要な要因であることが確認されており、この2つの環境要因の低下は問題である。

国家教育省が毎年全国統一で実施する学力診断テスト(卒業試験)の結果(10段階評定)を調べてみると、若干の差異はあるが両グループとも分権化後に学力が低下している。予算の多いグループの平均点は、地方分権化の実施前(1999/2000年度)は5.8点、実施直後(2002/2003年度)は5.4点であり、予算の少ないグループの平均点は各々5.7点及び5.2点である。中退率に関しては、予算の多いグループは1999/2000年度の2%から2002/2003年度の1%と減少したのに対し、予算の

少ないグループでは1%から2%と増大している。

中央から地方への権限が委譲されても、地方政府によって、住民の社会的厚生が改善がなされなければ、地方分権化の意図は達成されていない。本研究では、分権化後の予算の学校間格差の拡大に着目し、学校予算を起点とした因果モデルによって新たな取組みの必要性を探った。教育環境の水準の低下や一定の学校投入から生じる効果の減少は、予算の少ない学校により顕著に現われる傾向があるが、これは従来からインドネシアの教育課題であった学校間格差の問題を更に深刻にする可能性がある。教育環境における格差は正に向けて地方政府に期待される役割は大きい。

SBM及びCBEの導入により、分権化後は学校及び地域の教育リーダーとしての校長の役割は重要性を増すと考えられる。分析2の結果からも校長資質から教員資質への影響力は分権化後も2つのグループ両方において比較的安定しており、重要性が示唆されている。また、学習環境への新たな有意性も現れている。このことから、教育成果の向上を促進する施策のひとつとして、分権化前はほとんど行われていなかった校長研修等を実施し、経営能力やリーダーシップなど校長の資質向上をはかる取組みが望ましい。

教育環境の水準比較で明らかになった予算の少ないグループにおける分権化後の学校予算水準の低下を考慮すると、該当する学校への補助金等の供与を通して、予算の安定的確保を助け、学校間格差の是正をはかることが重要であろう。但し、予算の少ないグループにおいて、学校予算は教員資質を介して間接的に教育成果を上げていくことが、モデルの妥当性の検証を通して明らかになっている。影響力は低下するものの、分権化後もその因果関係の有意性は高い。このことから、資金のみを供給するのではなく、教員資質の向上をはかる取組みを合わせて行うことが重要と考えられる。

教員資質の改善が教育成果の向上に重要な役割を担うことは、予算の多いグループ、少ないグループとも共通しており、教員研修の実施による資質向上の取組みは、予算の少ないグループを中心に広く行われることが望ましい。しかし、教育環境水準の比較分析から、予算の少ないグループ

において分権化後、教員資質の水準にも低下が認められたことを考慮すれば、更に重点的な取組みも必要とされよう。因果モデルの分析の結果から、経験豊かで等級も高い教員が常勤として多くいることが教育成果の向上に有効であることが示唆されている。分権化後は中学校の教員人事の権限は県・市の教育局に委譲されている。問題を抱える学校に対して、教員配属上の工夫が行政的に取組まれることも、教育の質の均等化のためには重要と考えられる。

5. おわりに

本研究では、教育環境と教育成果に関する因果モデルの検証を通じ、インドネシアの地方分権化がSLTP公立校の教育活動に及ぼした影響を指摘するとともに、いくつかの課題について提起することができたと考える。

最後に本研究の制約と今後の研究の方向性について触れておきたい。

- (1) 地方分権化の過渡期における分析であり、今後どのように変化していくか、分権化の推移を見ながら継続的に観察していく必要がある。
- (2) SLTP公立校よりも教育環境において一般的に劣るマドラサやSLTP私立校を分析から除外し、且つ予算状況に基づきグループ分けを行ったことにより、所属する生徒の属性はある程度コントロールされていると考えられるが、生徒に関する社会経済指標や地域性等のデータを併せ用いることによって、さらにきめ細やかな分析が可能となるであろう。

以上の2点を今後の研究課題としたい。

参考文献

- 江原裕美 (2005) 「ブラジル初等教育改革における分権化と学校自立性の強化」『帝京大学外国語外国文学論集』第11号、pp.56-91
- 育藤泰雄 (2004) 「教育の市場化・民営化の行方 南米チリ20年間の経験」『国立教育政策研究所紀要』第133集、pp.7-19
- 田中美由紀・山崎瑞紀・柳井晴夫・鈴木規夫 (1998) 「個人の資質と大学の専門分野での適応に関する因果モデルの検討」『教育心理学研究』46号、pp.262-270
- 田部井明美 (2001) 『SPSS完全活用法 共分散構造分析によるアンケート処理』東京図書、p.141
- 戸井敦子・牟田博光 (2003) 「インドネシア中学校建設プロジェクトのインパクト評価」『日本評価学会第3回全国大会報告論文集』日本評価学会、pp.37-41
- 豊田秀樹 (1998) 『共分散構造分析「入門編」構造方程式モデリング』朝倉書店
- 中矢礼美 (2005) 「インドネシアにおける自律的学校経営に関する考察」『国際教育協力論集』第8巻、第2号、pp.51-62
- 中矢礼美 (2001) 「インドネシアの学校教育における地方分権化に関する研究(2)」『教育学研究紀要』第47巻、第1部、pp.384-389
- 中矢礼美・諏訪英広 (2001) 「インドネシアの教育行政 - 地方分権化政策分析に向けての予備的考察」『徳島文理大学研究紀要』第61号、pp.123-129
- 松井和久 (2003) 『インドネシアの分権化』アジア経済研究所
- 百瀬侑子 (2001) 『「教育」を通して知るインドネシア インドネシアの教育制度・教育事情』研究工房ブスタカロカ、p.9
- 森岡修一 (2000) 「ペレストロイカ期における教育改革」、日本教育経営学会編『諸外国の教育改革と教育経営』、玉川大学出版社、pp.61-74
- 山本嘉一郎・小野寺孝義 (1999) 『Amosによる共分散構造分析と解析事例 第2版』ナカニシヤ出版
- Bowls and Gintis (1976). *Schooling in Capitalist America*, Basic Books.
- Case and Dealton (1999) *School Input and Educational Outcomes in South Africa*, *Quarterly Journal of Economics*, Vol.114, No.3, pp.1047-1084.
- Coleman et al (1966). *Equality of Educational Opportunity*, U.S. Government Printing Office.
- Department Pendidikan Nasional (2003). *Indonesia Educational Statistics in Brief*.
- Edmonds, R.R. (1979). *Effective School for the Urban Poor*, *Educational Leadership*, Vol. 37, No.1, pp.15-27.
- Fuller, Bruce (1987). *What School Factors Raise*

- Achievement in the Third World?, *Review of Educational Research*, Vol. 57, No.3, pp. 255-292.
- JICA (2005). *The Study on Regional Educational Development and Improvement Program in The Republic of Indonesia*, IDCJ PADECO Co, Ltd., p.14.
- Jimenez, E & Paqueo, V (1996). Do Local Contributions Affect the Efficiency of Public Primary School?, *Economics of Education Review*, Vol. 15, No.4, pp. 377 - 386.
- Ministry of National Education (2003). *Indonesia Education Statistic in Brief*.
- Psacharopoulos and Woodhall (1985). *Education for Development*, World Bank.
- (2006.1.30受理)

The Influence of the Decentralization in the Area of Junior Secondary Education in Indonesia An Empirical Study Using Causality Model on the Relationship between Educational Environment and Outcome

Atsuko Toi Hiromitsu Muta

Tokyo Institute of Technology
ato@hd.hum.titech.ac.jp muta@hum.titech.ac.jp

Abstract

While there is a trend towards decentralization in educational policies, it is not a panacea for the problems faced by developing countries. The disparity between regions or schools in educational achievements has been pointed out as one of the defects attributable to decentralization.

This paper offers a justification for such proposition by studying the case in Indonesia. The method of covariance structure analysis was used in order to find out the influences and effects of the factors related to educational environment upon educational outcome, and to make comparison between before and after the introduction of the decentralized system in Indonesia.

The main findings and conclusion are following; generally speaking, the increase in the school budget improved the quality of educational environment, and led to higher test scores and lower dropping out rates. Such positive results, however, turned out to be less significant, and the relationship between factors became weaker during the period after the decentralization system had started. This was particularly the case for the group of schools whose budget level was relatively small.

In addition to the efforts to improve the quality of teachers and headmasters, it is important to consider the measures aimed at correcting the disparity between schools as well as regions.

Keywords

Indonesia, decentralization, junior secondary education

【研究論文】

構造調整借款20年間のレビューからみる日本政府の政策と判断

田中 弥生

東京大学

yayoi-tanaka@civil.t.u-tokyo.ac.jp

要 約

外務省はODA評価の一貫として日本の構造調整借款について初めて包括的なレビューを行った。1986年、構造調整借款の開始以来、約20年間で9,767億円が投じられている。本論では開始時期の政策背景を探ろうとした。そこでみえてきたのは、日米貿易摩擦に取り組み日本政府の葛藤と資金環流措置であった。当時、構造調整借款は総額650億ドルの資金環流措置のパッケージのひとつで、足の速い援助モダリティであることから環流目的に適しているとして選択されたのである。そこで問われるのは、日米貿易摩擦解消のための資金還流という目的を背負いながらもODAとして構造調整借款を選択したことの妥当性であろう。また、ODA政策を超えて政治的判断に直結した案件にかかわる政策評価の対象をどの範囲までに及んで考えるべきなのか。本論は、政策評価の課題についても問題提起した。

キーワード

構造調整借款、日米貿易摩擦、前川レポート、資金環流措置、政策評価

はじめに：本研究の意義

本研究は日本の構造調整借款のレビュー結果（外務省）をベースに、大きく2点を明らかにしようとした。ひとつは構造調整借款開始に到る日本の政策的背景と判断を明らかにすることである。構造調整借款は、世界的な規模で起こった債務危機問題に対応すべく、1980年より、新しい援助モダリティとして世界銀行で始められたもので、80年代の援助方法に影響を与えている。日本は、1986年より構造調整借款に着手するが、債務危機問題発生の時期、あるいは世界銀行が構造調整借款開始した時から数年のギャップがある。この間

の日本の動向を探ることによって、構造調整に着手するに到った政策判断の内容を明らかにしようとした。

もうひとつは、評価のあり方に関するものである。構造調整借款の背景を探ると、そこにはODAの枠組みを超え、日米貿易摩擦解消や資金環流措置という政治的な判断に基づく政策目的の存在が見えてきた。すると政治的な判断に直結した案件の評価対象をどの範囲まで及んで考えるべきなのかという疑問が浮上した。本論ではこの点についても問題提起した。

1. 外務省「日本の構造調整借款レビュー」からみる20年の推移と目的

(1) 調査目的

外務省は平成16年度ODA評価の一貫として、日本の構造調整借款のレビューを実施し、筆者は本レビューの評価主任をつとめた。昨今、英国、世界銀行などが実施している一般財政支援という援助モダリティは国際的な援助潮流になりつつある。日本でも財政支援への取り組みが検討されている。そこで、財政支援と同様にプログラム型（ノンプロジェクト型）の援助として位置付けられる構造調整借款をレビューし、その課題を明確にすることによって、今後、財政支援取り組みのための示唆を引出すことを本レビューの目的としている。

(2) レビュー対象設定の困難性

プログラム型援助には複数の種類があり、その類型も援助機関によって異なっていた。例えば、世界銀行の場合には、調整貸付の中に4種類のスキームに分類されている¹。日本の外務省ではノンプロジェクト型は、商品借款、構造調整借款（SAL）、セクター調整借款（SECAL）、セクター・プログラム・ローン（SPL）に区分されている（外務省 2002a,p.125）。世界銀行と異なり、日本では調整融資というカテゴリーは一般的ではなく、構造調整借款という言葉がしばしば用いられているが、その概念はあまり明確になっていなかった。日本の構造調整借款の多くは世界銀行との協調融資で行われており、世界銀行と同様に被援助国の制度改革を目的として実施される融資を調整融資であると捉え、構造調整借款（SAL）とセクター調整借款（SECAL）をレビューの対象とした。

対象の定義をしておく。まず、構造調整借款（SAL）であるが、日本では「借入国の総合的な経済政策改善と制度改革を支援するための借款であると定義されており、輸入商品の購入を借款の対象として、必要に応じて構造調整計画実施のためのコンサルタント・サービスを行う」としてい

る（外務省 2002、p.130）。具体的には、借入国と援助側がマクロ経済安定のために必要な政策や政策改変について話し合い、合意がとれたものについて構造調整借款供与のコンディショナリティとする。融資対象は借入国の輸入決済であり、国際収支支援の意味合いも有する。

次に、セクター調整借款（SECAL）であるが、「特定セクターの総合的な政策改善と制度改革を支援することを目的とする」（外務省2002、pp.130-131）ものであると定義され、輸入商品を借款の対象商品、あるいは当該セクターの開発に必要な輸入資機材である。また、当該セクターの制度改革に必要なコンサルテーションも行ってゆく。世界銀行は構造調整借款とセクター調整借款は相互補完するものであるとしているが（World Bank 2001）、国全体の経済政策の改変と運営を実行できるだけの能力が借入国に不足している場合には、SECALが用いられており、80年代、80年代を通じて両者は共存している²。

レビューの対象期間は、日本の構造調整借款の始まった1986年から最近の融資が行われた2002年度までである。

(3) 評価可能性の問題

この調査では評価ではなくレビューと行うとしている。レビューとした背景には評価可能性の問題があった。ひとつは、構造調整借款の形態が細分化され、その区分体系が援助機関によってまちまちであったために評価対象を確定することが困難であったこと、そして20年前に遡りデータを収集することも困難であり、評価可能性が低いと判断したのである。

(4) レビュー結果からみる構造調整借款の全体傾向と融資目的

レビューによって、複数の発見事項があったが、ここでは全体的な推移と構造調整借款の目的について着目する。

全体的な実績の推移と背景

図1は日本の構造調整借款20年間にわたる供与

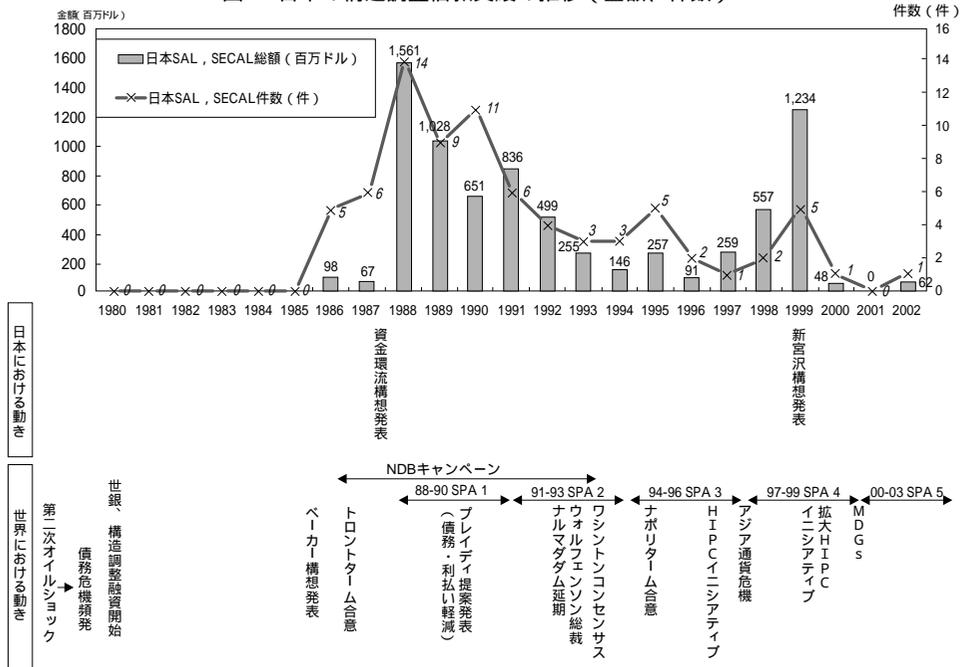
額の増減を示したものである。構造調整借款は、第2次石油危機による石油価格の高騰、米国のインフレ抑制政策により実質金利が上昇したことにより、危機に陥った開発途上国を支援するために1980年、世界銀行によって開始された。世界銀行の調整融資は他ドナーと協調融資の形で実施されており、日本は最大のパートナーとなっている。日本は、1986年から2002年までに、78件、総額9,767億円の構造調整借款を承諾している。年度別の傾向をみると、案件数、承諾額ともに1988年にピークを迎えている。その後、1992年から減少を続ける。1997年頃から再び承諾額が増加し1999年にピークを迎えている。

構造調整借款承諾額の増減の背景には何があったのかを説明しようとしたが、表の下に「日本における動き」「世界における動き」の欄を設け、前者には関連しそうな日本の政策、後者には、世界銀行や日本の構造調整借款に関連すると想定された世界の出来事を記していった。これらは、関係機関からのヒアリングと資料で確認されたものを記している。

1988年のピーク時の背景として、資金環流措置とベーカー構想を挙げた。当時は開発途上国にとって、プロジェクトを推進するよりも国際収支の悪化を乗り切る事が急務であり、ノンプロジェクト援助が援助潮流の主流となっていた時代でもある。また、86年、87年の融資対象地域はアフリカであるが、当時、アフリカは国際収支の悪化に直面しており、世界銀行、IMFなど複数ドナーは、アフリカの経済構造改革プログラムの資金確保、援助有効性を高めるための協調機会をつくるためにSPA(Special Program of Assistance for Africa)を設立した。日本はSPA開始当初、積極的な姿勢を示し1987年には6件、9億円、1988年には9件、69億円を承諾している(外務省2005)。

1999年には第2のピークを示している。1997年以降、構造調整借款の実績が再び伸びているが、その殆どは東南アジアを対象としたものであり、97年に始まったアジア通貨危機への対応である。アジア通貨危機後、国際収支のバランスを取り戻すための短期的資金需要と同時に金融システムなどマクロ経済の安定化のための中長期の資金需要

図1 日本の構造調整借款実績の推移(金額、件数)



(出所) 外務省「調整融資のレビュー報告書」(2005) p.10

の双方が高まった。日本は「新宮澤構想」などを通して短期的および中長期の支援を行ったが、主として構造調整借款を用いて行われた（外務省2002b）。

1992年から1996年まで構造調整借款の実績は下降を続けている。この間の「日本の動き」については、確認できる情報やデータがみつからなかったために記していない。「世界における動き」では、世界銀行やIMFによる調整融資などへの批判が高まった時期であること、HIPCイニシャティブなど重債務国への債務削減を記しているが、これら日本の調整融資にも影響を与えているのではないかと示唆した。

では、ODA全体にとって、構造調整借款はどのような意味があったのか。1986年から2002年までのODA総額に占める構造調整借款の割合は全体の5%前後と推定される³。だが、構造調整借款が開始された1986年は、第3次中期目標（1986-1992）の時期に当たるが、当時のODA実績目標は400億ドルである。同時期の構造調整借款は46.4億ドルであり、同時期のODA総額（実績目標）に占める割合は13～14%である。したがって、この時期における構造調整借款のプレゼンスは相対的に大きく、これまでプロジェクト中心であった日本のODAに影響をもたらしていた時期と考えられるだろう。

実績からみる構造調整借款の目的

構造調整借款の実績の推移をみると、時代により実績額の増幅に違いがあり、構造調整を用いた目的は、時代により異なっていたのではないかと予想される。そこで、レビュー結果の・援助対象

地域、目的およびコンディショナリティの推移に着目する。

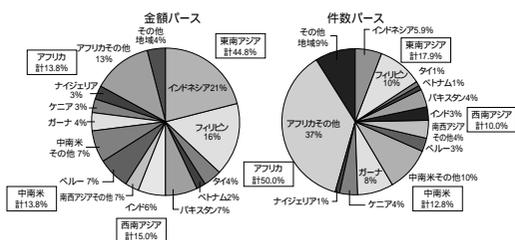
まず、図2に示した構造調整借款の実績を対象地域別にみると、件数ベースではアフリカが全体の50%、東南アジアが17.9%、承諾額ベースでは、東南アジアが全体の44.8%、アフリカが22.6%を占めている。また、各地域とも少数の国に融資が集中していることもわかる。ODA全体に占めるアフリカ援助の比率は10%前後であることに比較して、アフリカ向け融資が多いことが特徴である。

さらに、図3のアフリカ向け借款の内訳に着目すると、1986年は16億円、1987年は9億円、1988年には69億円、1989年は11億円と、実績額の振幅が大きい。同時に開始直後数年で相当額の構造調整借款が提供されている。

1980年代のアフリカの状況に着目するとアフリカは自然災害、民族紛争、一次産品の輸出不振などから、公的機関からの借入が増加し債務問題が深刻化していた。このような状況下、1985年、世界銀行のイニシャティブによって、アフリカの国際収支問題に取り組むべく、ドナー援助協調の枠組SPA（Special Program of Assistance for Africa）が作られた。日本もSPAに積極的に参加したが、86年の最初の構造調整借款はSPAに参加するかたちで開始された（外務省2003）。

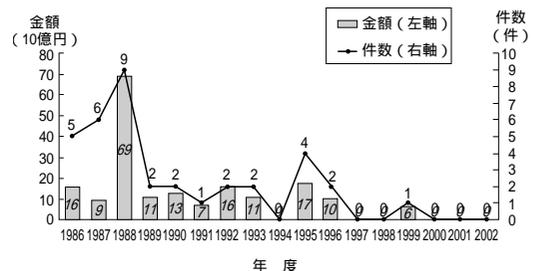
次に東南アジアであるが、主たる対象国はインドネシアとフィリピンである。インドネシアへの借款は87年、88年の2年間、フィリピンへの借款は88年から90年の3年間に留まり、その後、アジア通貨危機対応として融資される98年まで融資はされていない。構造調整借款融資が開始された当時の両国の状況は、円高ドル安の影響を受け、86

図2 日本の構造調整借款の対象国割合



(出所) 外務省「調整融資のレビュー報告書」(2005) p.23

図3 対アフリカ構造調整借款の実績と推移

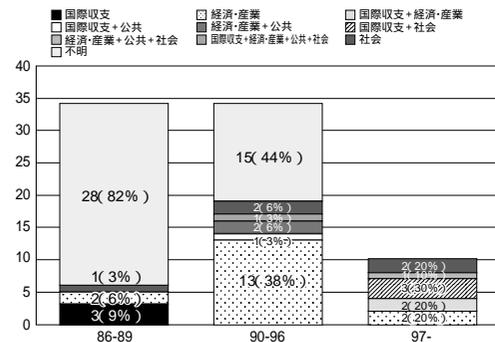


(出所) 外務省「調整融資のレビュー報告書」(2005) p.28

年にかけて対外債務比率（対GDP比）が急上昇した時期であり（インドネシアは約70%、フィリピンは95%）国際収支支援が必要な状況にあった⁴。

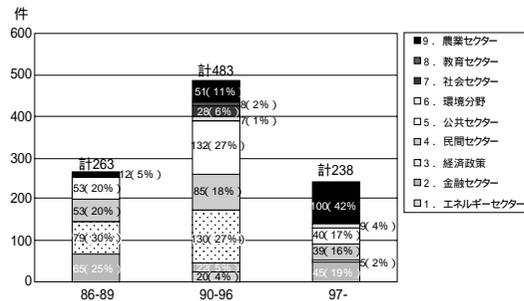
次に構造調整借款の目的とコンディショナリティに着目してみよう。80年代は国際収支の改善、経済改革をその目的の中心としていたことがわかる。図4は、目的別構造調整借款の推移を示したものである。不明が多く、データとしての限界は否めないが、86-89年代は国際収支支援と経済・産業、90-96年は経済・産業、97年以降は国際収支をベースに、経済、産業、社会、公共と複数目的がまんべんなく組み合わせられている。図5はコンディショナリティについて説明したものだが、金融、経済政策は86年開始から一貫して対象となっているが、その上で、公共、社会、農業などのセクターなどの組み合わせの幅が広がっている。

図4 目的別構造調整借款件数の推移



(出所) 外務省「調整融資のレビュー報告書」(2005) p19

図5 セクター別コンディショナリティ推移



出所：外務省「調整融資のレビュー報告書」(2005) p.20

目的別、コンディショナリティから、構造調整借款の目的は、対象国の国際収支・経済政策の改善をベースにしなが、次第に社会開発やガバナンスの改善をも対象にしてきていることがわかる。

しかしながら、日本は世界銀行との協調融資で構造調整借款を行っていたため、これらの目的やコンディショナリティは世界銀行の意向をも反映している。したがって、これらをもって、日本が構造調整借款を選択した際の目的とは言い切れない。ここで明らかに言えることは、日本は世界銀行などの援助潮流に合わせて構造調整借款を行っていたということである。

2. 問題提起

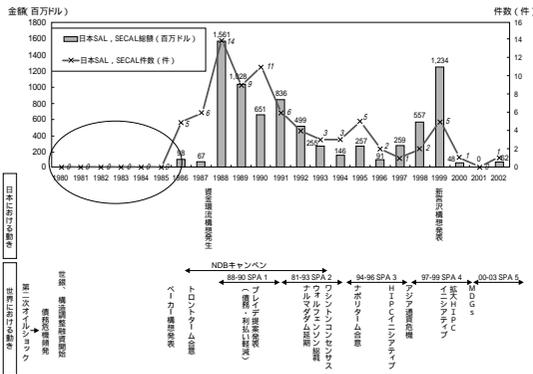
(1) 構造調整借款と日本の目的

では、日本は何故、構造調整借款を始めたのか。援助の潮流は理由のひとつであるが、それだけが理由ではないだろう。構造調整借款を開始した当時、政策対話のあり方やアカウンタビリティのあり方について疑問視する意見も出ていた⁵。

また、世界銀行が構造調整借款を開始したのは1980年、債務危機がメキシコで起こったのは1982年である。アフリカの債務問題も80年代初頭には顕著になっていた。ところが、日本が構造調整借款に最初に着手したのは1986年で対象地域はアフリカである。債務危機問題が顕著であったラテンアメリカには1988年に初めて構造調整借款を出している。この間の数年の時間差は何を示唆しているのだろうか。図1の日本の構造調整借款の実績の推移では、「日本の動き」として「資金環流措置」を挙げている。これが、具体的に構造調整借款とどのような関係にあるのか。

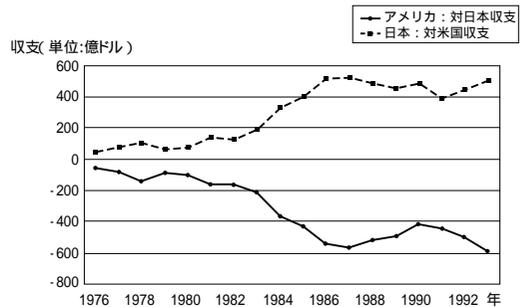
そこで、図6の赤字で囲んだ構造調整借款開始前後数年の日本の動き、資金環流措置との関係を探ることによって、構造調整借款という援助モデルを選択した日本の目的を明らかにすることができるのではないかと考えた。

図6 本論の着目点：構造調整借款着手前後
1980年から数年間のプランク



(出所) 外務省「調整融資のレビュー報告書」(2005)p.10に加筆

図7 アメリカと日本の貿易収支



(出所) 世界平和研究所「中曽根内閣史 - 理念と政策」(1995) p.581

(2) アプローチ

20年前の情報収集は、ヒアリングとデータ収集調査によって行った。ヒアリング対象者は、資金環流措置および資金環流計画に直接携わった大蔵省国際金融局局長、債務危機発生後に銀行諮問団 (Bank Advisory Committee) のメンバーとして他国の民間銀行、IMF、債務国との交渉に立ち会った元東京銀行常務、構造調整借款実施に関与した元経済企画庁職員、元海外経済協力基金職員、国際機関研究者など有識者である。

また、ヒアリング結果にあわせ、それを裏付けるデータを収集、分析した。また、調査対象は、ODAによる構造調整借款に限定せず、民間銀行や輸出入銀行などの動向も探ることにした。

3. 構造調整借款の開始前後の背景

まず、日本を取り巻く国際情勢および、国際的に勃発した債務危機問題の状況について触れ、次にこのような国内・国際状況の中で日本政府がどのような判断を下し策を打っていったのか説明する。

(1) 80年代の日本を取り巻く情勢

日米貿易摩擦

この時期、日米間の貿易不均衡は急増し、貿易摩擦が日米関係の阻害要因となりはじめていた。

米国は1980年代初頭の高金利のため極端なドル高になり、同時に財政赤字が急増、その結果、貿易赤字が急増していった。1985年9月ドル高是正のための通貨調整合意、すなわち「プラザ合意」が、先進5カ国の間で行われた。これによって、円高ドル安へと動いていった。しかしながら日本の貿易黒字は増加を続け、米国の対日貿易摩擦は解消されなかった。80年の対日貿易赤字は104億ドルであったのに対し86年には544億ドルになっている(図7)。

一方、日本は国際競争力を増し、景気は緩やかな回復に留まっており、輸出は伸びたが、輸入は増えなかった。その結果、貿易黒字、とくに対米黒字が急増し、80年には69億ドルであったのが、86年には514億ドルになっている。経常黒字の対GNP比は81年にはゼロであったのが、86年には4.3%に到達しており、日本政府は「危機的状況」とこの状態を評している(世界平和研究所1997)。この当時、米国では議会を中心に、日本は不公正な貿易相手国であるというパーセプションが広がり、対日批判が米国民に広がっていった。

日米間の貿易摩擦を憂慮した中曽根内閣は、輸出の自主規制や個別産業分野の市場解放措置を実施していった。輸入数量や価格規制の対象となったのは繊維、鉄工、自動車、半導体などであった。次に、問題の焦点になったのは日本市場の閉鎖性であり、エレクトロニクス、電気通信、医療品など市場個別分野において日本が自主的に輸入促進を行おうとした。具体的には首相自ら国民に外国製品輸入を呼び掛け、外国大使館から関係者を徴

集し不満を聞く会を設けたりした。

前川レポート

しかしながら、個別分野で市場解放措置を実施しても日米貿易不均衡は容易に解消されなかった。結局は、「輸出が増えやすく、輸入が減り難い経済構造」を変えるしか貿易不均衡問題を解決する抜本策はない（世界平和研究所1997）と、中曽根首相は判断する。1985年、中曽根首相は前日銀総裁 前川春雄氏を座長とする「国際協調のための経済構造調整研究会」を発足し、翌年4月には、報告書「前川レポート」が首相宛に提出された。内容は内需拡大を柱にするものであったが、国際通貨危機の安定化と金融の自由化・国際化への対応策が謳われ、累積債務問題への対応として、開発途上国への公的資金フローの拡充、国際開発機関の資金基盤の強化およびその機能のいっそうの効率化に努めることが記されている。また、経済・技術協力の推進の項目では、政府開発援助（ODA）の拡充、民間援助団体（NGO）の活用的重要性、アンタイド化の推進などの必要性が謳われていた。

このように当時の内閣は、1980年代に顕著になった日米貿易摩擦問題に対して抜本的な経済構造の改革が必要であると判断しているが、その対応策の一貫として、途上国への援助の拡充や公的資金のフローの拡充などを挙げている点は注目に値する。

（2）国際社会情勢：債務危機問題

1980年代は、債務危機問題が勃発し世界経済に緊張をもたらした時期でもある。そこで、債務危機の状況や国際社会の対応状況についてみてゆく。

インフレ抑制政策の負の影響

1979年、米国経済衰退を背景に、米国連邦準備理事会（FRB）は金融政策をインフレ抑制最優先に転換した。その結果、インフレ率は低下し、先進国および国際金融市場での実質金利が急上昇し、同時に先進諸国の経済成長率が低下することになった。また一次産品価格は低落し、中低所得

国の輸出は減少した。

この時期に累積債務問題が中・低所得国を中心に勃発するが、中でも中所得国に問題が集中していた。世界の債務残高における中所得国のシェアは70年代、71%であったものが、82年には86%へと上昇している（平田 1993）。

地域的にはラテンアメリカに顕著であり、1982年メキシコが対外送金を停止し、債務危機問題が顕在化していった。ラテンアメリカの債務問題の特徴は民間資金からの債務が大きいのが特徴である。

ラテンアメリカに次いで債務問題が深刻化していたのはアフリカ諸国である。石油ショックなどの影響による一次産品の輸出不振に加え、自然災害、民族紛争などの打撃を受け債務は70年代から80年代にかけて8倍以上に膨れ上がっている。1970年代には55億ドルであったのが、1980年代には440.9億ドルに増え、その内容は400.6億ドルが公的部門からの債務、43億ドルが民間部門からの債務となっている（毛利 1995）。

債務危機発生メカニズム

では、どのように債務危機は起こっていったのか。1973年第1次石油危機により、石油価格は高騰し、輸入石油に依存していた先進諸国の多くはスタグフレーションと貿易収支の悪化、財政赤字のダメージを受け、対外援助力を失っていった。

一方、大手の民間銀行は先進諸国に貸付先を見出しにくくなり、原油価格高騰により潤った産油国の資金（オイルマネー）に着目した。銀行はこれらの金を集め、中低所得国に貸付先を求めていった（ボルガー、行天、1992）。

原油価格と機械などの完成品価格の高騰、輸出一次産品の数量不振と価格の下落、先進諸国不況にともなう出稼ぎ送金の減少など、中低所得国を襲ったこれらの出来事によって、外貨獲得はより困難になっていった。そこで、公的機関に比較し融資条件が緩やかな民間銀行からの借入が行われるようになったのである。当時高インフレ状況にあったために、金利が相対的に低かったことも借入を促す要因になっていた。民間銀行は、比較的リスクの少ない中所得国を中心に貸出しを行って

表1 対外債務の債権者別構成:中所得国と低所得国(%)

	1970	1974	1978	1982
	中所得国グループ			
DOD合計	100.0	100.0	100.0	100.0
公的資金	59.3	50.2	41.3	38.3
多 国 間	14.6	12.8	12.3	14.0
2 国 間	44.7	37.4	28.9	24.3
民間資金	40.7	49.8	58.7	61.7
輸出信用	15.8	12.5	9.0	5.3
金融市場	22.2	35.3	49.0	56.2
	低所得国グループ			
DOD合計	100.0	100.0	100.0	100.0
公的資金	89.6	88.2	86.7	87.2
多 国 間	18.2	20.1	26.8	34.6
2 国 間	71.3	68.1	59.9	52.5
民間資金	10.4	11.8	13.3	12.8
輸出信用	7.4	6.8	5.6	3.3
金融市場	2.0	4.9	7.7	9.5

(出所) 平田喜彦「現代国際金融の構図」(1993) p.282、
World Bank「World Debt Tables」(1983-84 edition)

いた⁷。第1次石油危機から第2次石油危機の間、途上国向けの民間銀行からの貸付は5倍に膨れ上がっている⁸。借入国の債務は増加し、新規融資は銀行への利払いに充当される状態になっていた。

表1に着目すると70年代から90年代に到る中所得国の債務構成は、民間銀行の比率が40.7%から61.7%、公的資金は59.3%から38.3%であり、債務問題の原因は民間銀行に対する債務の急増であるといつてよいだろう。

そして、1979年のFRBのインフレ抑制政策によって、実質金利は上がり民間銀行は貸出しを渋り始め、債務国側は利払いが不可能な状態に陥っていった⁹。

他方、アフリカなどの低所得国で起こった債務問題はラテンアメリカなどの中所得国で起こった債務問題とその内容を異にしている点がある。それは債務の構成であり、公的債務比率が高いことである。400.6億ドルの公的債務に対して民間部門からの債務は43億ドルであるので、公的債務は民間債務の9倍以上である。この点は、後に説明する日本の資金還流措置のパッケージの中で輸出入銀行と海外経済協力基金との間の役割分担のあり方に影響してくる点である。

(3) 国際社会の債務危機対応

銀行諮問団(BAC)の結成

ラテンアメリカ地域を中心起こった債務危機に対して米国政府は即座に対応を始めた。1982年8月メキシコ危機が起こって1週間も経ないうちに、IMFに対しては拡大信用供与39億ドル、公的債務のリスケジュール、先進国中央銀行によるつなぎ融資などを交渉した。また、Bank Advisory Committee(BAC)という、米国、日本などラテンアメリカ向け債権銀行から構成される銀行諮問団が同年8月に結成され、リスケジュールや新規融資を取りまとめ、債務国対応が進められた¹⁰。ここでは、IMFが債務解決に向けて経済構造改革のイニシアティブをとったが、債権者である銀行と債務国の交渉は両者間で行われ、そこに世界銀行、米国財務省がオブザーブ参加していた。

ベーカー構想

1985年、レーガン政権下、ベーカー構想が発表される。累積債務国45カ国への対応を目的としたもので、まず、IMF・世界銀行を中心に債務国との協議の上、経済構造調整プログラムを実施する。他方で、各国民間銀行は3年間で、総額200億ドルの新規融資を行うという構想である。この構想の背景には、債務危機に直面した国々の問題は資金の流動性にあるので、それを高めるためにはニューマネーが必要で、したがって新規融資を促す必要があるという考え方であった。先進国政府、IMF、世界銀行などの国際機関は、触媒機能は果たすが、民間銀行が債権回収をするための肩代わりには手を貸さないというスタンスをとっていた(幸島2002, p.128)。民間銀行の過剰貸し付けによって引き起こされた債務問題は、民間銀行によって解決するべきであるという、基本的なスタンスがみてとれる。

しかしながら、民間銀行からの新規融資が期待とおりに集まらずベーカー構想は頭打ちになっていた¹¹。

ブレイディ構想

1989年、ブッシュ政権下、ベーカー構想に代わるものとして、新債務戦略(ブレイディ構想)が

発表された。ここでは、IMFが債務国との間で政策対話を行い、構造調整借款協定を結ぶ。IMFとの協定が成立すると、民間銀行は、債権の債券化（市場売却が可能）が認められる一方で、債務元本の一定割合削減、利払いの一定割合削減、新規融資のいずれかを行う。世界銀行とIMFは債務国の信用保証を行うが、先進国政府はこれらの国際機関に対して融資を行うというものである（海外経済協力基金史2003）。プレイディ構想の特徴は、債務元本を現在価値に割り戻すところにあり、実質的な元本削減を正当化する理由にもなっている（幸島2002）。ペーカー構想では債務危機は資金の流動性の問題であるとしていたが、プレイディ構想では債務支払い能力に問題があると考えてられていた。

（4）日本政府の政策判断

1986年、中曽根首相は深刻化する日米貿易摩擦を憂慮し、当時、大蔵省国際金融局長であった内海氏に資金環流の策を講じるように、指示をする。丁度、「国際協調のための経済構造調査研究会」より報告書（前川レポート）が首相に提出された年である。大蔵省は2条件を提示して、資金環流計画を策定した。2条件とは、ひとつは、アンタイドであること、もうひとつは黒字の大部分は民間に帰属しているので、民間資金の活用を促すことであった。

資金環流措置とペーカー構想

1986年ワシントンのIMF・世界銀行総会で、日本政府は100億ドルの資金環流計画を発表した。これは前年に提出されたペーカー構想に対応するもので、内訳は、36億ドルが日本輸出入銀行からIMFへの貸付、アジア開発銀行基金ファンド財源補充（39億ドル）、世界銀行ジャパン・ファンド創設（20億ドル）であった。87年、日米首脳会談で、中曽根首相は更に200億ドルを資金環流にまわすことを表明し、300億ドルの資金環流計画となった。その後、87年に発足した竹下内閣に受け継がれ89年まで推進されることになった。89年のアルシュ・サミットにおいて、新たに350億ドルが拡充され、総計650億ドルの資金環流計画とな

った（海外経済協力基金史2003）。資金環流措置のポイントは、アンタイドと民間資金の環流である。仮に公的資金を投じたとしても、それが民間資金の環流を何倍かに増やす触媒機能を果たすことが求められた。そこで、考え出されたのが以下2つの方法であった。

第1に、世界銀行などの国際機関に「ジャパン・ファンド」を創設することである。ここで、プロジェクト開発を行うが、それによって資金需要が増えることになる。国際機関は円建てで資金調達を行う。これによってジャパン・ファンドとして積んだ基金の複数倍の資金環流が実現することが期待されていた。

第2に、日本輸出入銀行、海外経済協力基金の資金を動員することである。資金環流目的のためには、これらからの資金動員はアンタイドかつ「足の速い」方法で行われる必要がある。したがって、案件形成に時間を必要とし、支払いも長期間にわたって複数に分けて行われるプロジェクト型の支援は、資金環流目的には適さない。そこで、政策対話をベースに少ない支払い回数で実施される構造調整借款に白羽の矢があつたのである。表2は日本政府が発表した資金環流措置を海外経済協力基金が一覧にまとめたものである。この表からも、国際機関への拠出、輸出入銀行からの構造調整融資とODAの構造調整融資がパッケージになっていたことがわかる。

ちなみに、日本輸出入銀行のアンタイドローンの推移に着目すると（図8）、85年には1,010億円、86年には約3,380億円、87年には4,660億円と資金環流措置前後で、アンタイドローンの額が急増していることがわかる。

ラテンアメリカ債務危機対応

資金環流措置を進める中で、日本政府は、当時ラテンアメリカを中心に起こった債務危機問題に着目した。ラテンアメリカに着目したのは、その問題の深刻さばかりが理由ではない。ラテンアメリカの累積債務問題は米国国政にとっても重要課題であったからだ。日米貿易摩擦に悩む日本は、ラテンアメリカを日米共通の課題として捉えようとしたのである。

米国にとって、特に、メキシコ、ブラジル、ア

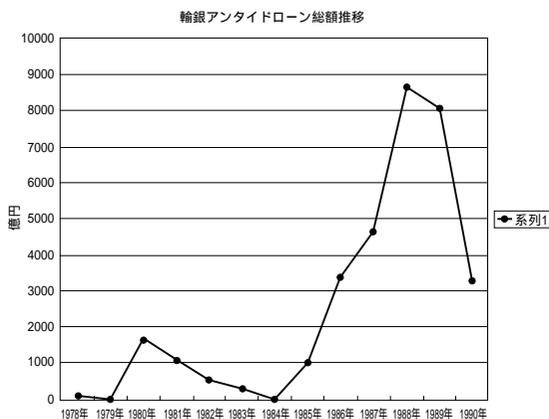
表2 資金還流措置とその拡充の概要

	300億ドル計画	650億ドル拡充計画		合計
		同左計画 実績見込み	拡充分	
1. 期間	1987年～90年 3年間			1987年～92年 5年間
2. 資金配分				
(1) 輸銀(民間銀行の協調融資を含む) 国際金融機関との協調融資 輸銀独自融資	90億ドル以上 (60億ドル) (30億ドル)	100億ドル	135億ドル程度 (新債務戦略適用国に対して約 80億ドル以上)	235億ドル程度
(2) OECF 国際金融機関との協調融資	30億ドル以上 (30億ドル)	55億ドル	70億ドル (新債務戦略適用国に対して約 20億ドル以上)	125億ドル程度
(3) 世銀等国際金融機関に対する出資 ・ 拠出その他 当初計画分(注) 追加分	180億ドル以上 (100億ドル) (80億ドル)	145億ドル	145億ドル程度	290億ドル程度
計	300億ドル以上	300億ドル	350億ドル程度 (新債務戦略適用国に対して約 100億ドル以上)	650億ドル程度

(注) 39億ドルがIDA増資とADF(アジア開発銀行のファンド)の財源補充のための拠出、36億ドルが外国為替特別会計からのIMFへの貸付、20億ドルが世銀での日本特別ファンドの創設

(出所) 国際協力銀行「海外経済協力基金史」(2003) p.215

図8 日本輸出入銀行アンタイトローン総額の推移



(出所) 日本輸出入銀行海外投資研究所「海外投資研究所報」(1992) 1978-1990より作成

ルゼンチンの債務国3カ国が重要であり、これらが一度に債務問題で破綻しないように、1カ国が倒れそうな時は、他2カ国は維持できるようにしようとした。そして、債務国の調整に応じて、ケース・バイ・ケースで救済しようとした。日本は、これらの3カ国を中心に、民間銀行と日本輸出入

銀行からの融資で救済しようとした。日本の銀行は、民間銀行総額200億ドルの新規融資を期待したベーカー構想の有力な協力者になった¹²。他国の銀行が貸出しを渋る中で、日本の銀行の貸出残高は1984年で290億ドルであったが、1986年には370億ドルにも上った(ボルガー、行天1992、p.326)。

ちなみに、輸出入銀行からのラテンアメリカ向け融資は、86年には4.56億ドルであったのが、87年には16.53億ドルと約3倍に急増している(海外投資研究所報1986,1987)

宮澤構想

ベーカー構想の核は、民間銀行による新規融資であり、日本の銀行は当初積極的に協力していたが次第に陰りがみえてきた。新規融資をしてもそれが米銀への利払いに充当されるだけであり、また、貸倒引当が国際水準に比較して低かったからである(幸島 2002)³。

そこで、1988年、日本政府はベーカー構想に代わるプランを考えた。これが宮澤構想と呼ばれるものである。宮澤構想のポイントは「三方一両損」

と呼ばれるもので、債務国、国際機関と各国の公的機関、民間銀行のそれぞれが負担をしようというものである。債務国がIMFと協議し、中期的な経済構造計画について合意し、補助金削減や増税など政治的リスクを負う。民間銀行は債務国との話し合いで、債務の一部を債券化して市場で売却する代わりに利払いについて軽減する。IMFは債務国の外貨準備を管理・保障し、債券化しないものには引当金を積む、というものであった。

債務国、国際機関と先進国政府、民間銀行のそれぞれが負担を担うという点が、宮澤構想の特徴であった。

宮澤構想はIMFには歓迎されたが、米国政府の支持は得られなかった。1988年は大統領選の年であり、ペーカー構想とは異なるプランが出てくることを良しと考えなかったようだ。つまり、米国は自国以外から異なるイニシアティブが出ることを好まず、従って日本からペーカー構想の限界に対応する別構想が出る事に同意できなかったのだという¹⁴。

ブレイディ構想への協力

1989年、ブレイディ構想が発表される。その直前に、ダララ財務次官補が来日し、ブレイディ構想への協力を要請している。内容は宮澤構想と酷似したもので、異なる点は元本削減が民間銀行向けメニューの中に含まれていた点である。

日本政府はブレイディ構想への協力を決める。最初のそれはメキシコ向けの融資である。1989年、民間銀行が有する債権は元本削減もしくは金利が減免され、メキシコ政府国債と交換される。メキシコ政府国債の担保には、米国財務省証券（ゼロ・クーポン債）が充当される。メキシコ政府が米国財務省証券を購入する際の資金は、IMF（17.3億ドル）、世界銀行（20.1億ドル）、日本輸出入銀行が（20.5億ドル）融資した。

ところで、ブレイディ構想のメディア関係者への事前説明は1989年3月10日午後11時に米国よりも数時間先に日本で発表されている。記事を出すのは米国でブレイディが発表してからという条件付きで行われた。しかし、その結果、ブレイディ構想の記事は米国よりも日本のメディアで先に紹介された。記事発表後はブレイディ構想が宮澤構

想と酷似していると話題になったという。ブレイディ構想のアイデアを宮澤構想が提供していたとしても、米国よりも先んじてイニシアティブを取らぬよう日本政府は配慮しなければならなかったという。だが、米国よりも数時間早い記事発表は日本側の密やかな自己主張だったのではないだろうか¹⁵。

以上、日本の構造調整借款が開始される前後数年の情勢と日本の政策判断についてみてきた。ここで明らかになったのは、日米貿易摩擦の対応策としての資金環流措置とその一貫として行われた債務危機問題など国際社会への対応である。ラテンアメリカで起こった債務危機対応に積極的に対応したのは、国際的な金融秩序の混乱を憂慮したこともあるが、同時に米国との関係を配慮したからで、米国と緊密な関係にあるラテンアメリカの問題を日米共通課題として捉えたのである。

では、ODAによる構造調整借款にとって、これらの政策判断は何を意味していたのだろうか。

4. ODA政策を超えた政策判断と構造調整借款

(1) 資金環流措置における構造調整借款の役割

構造調整借款の背景には、ODA政策を超えた首相による政策判断があった。当時、日本政府が最大の課題としていたのは日米貿易摩擦問題であり、その対応策としての資金環流措置が策定されている。資金環流の先として、債務問題に苦しむラテンアメリカやアフリカなどの地域が対象となっていた。ODAによる構造調整借款は、国際機関におけるジャパン・ファンドの創設、IMFへの貸し付け、輸出入銀行による融資から構成される資金環流措置のパッケージのひとつであった。ここで構造調整借款という援助モダリティが選ばれたのは、「足の速い援助」でかつアンタイドであることから、資金環流目的に適していたからである。そこで、資金環流措置におけるODA構造調整借款の役割と位置付けについて確認する。

資金環流措置

資金環流措置の対象期間である86年から92年の構造調整借款の合計金額は46.4億ドルである。資金環流措置予算総額は650億ドルであるので全体の7%であり、ODA構造調整借款が資金環流措置において果たした役割は決して大きいとは言えない。しかし、資金環流における海外経済協力基金分の予算（125億ドル）に占める構造調整借款の比率は43.8%であり、資金環流目的のODAにおける構造調整借款の役割は大きいといえよう¹⁶。

債務危機対応

では、ラテンアメリカの債務危機対応との関係はどうか。外務省レビューによれば、ODAのラテンアメリカ向けの最初の構造調整借款は1988年で7億円、翌年は12億円とさほど規模が大きくない。ではラテンアメリカ債務危機は何によって対応したのか。そこで、輸出入銀行によるラテンアメリカ向け構造調整借款とあわせて図9に示した。図9によれば輸出入銀行によるラテンアメリカ向け構造調整借款は例えば、1989年には輸銀は3,161億円で、ODA構造調整借款は21.58億円程度で、輸出入銀行による融資が圧倒的に多く、ラテンアメリカの債務危機対応には輸出入銀行の構造調整借款が対応していたと言えるだろう。

そして、ODAの構造調整借款に再度着目すると図2の地域別の実績が示すとおり、融資額では東南アジア、件数ではアフリカが高く、アフリカ

や東南アジアの債務問題や国際収支問題に対応していたといえるだろう。つまり、中南米などの中所得国の債務危機には輸出入銀行が、東南アジア、アフリカのそれについては基金が対応するという役割分担ができていたと思われる。

(2) ODAとして構造調整借款を選択することの妥当性

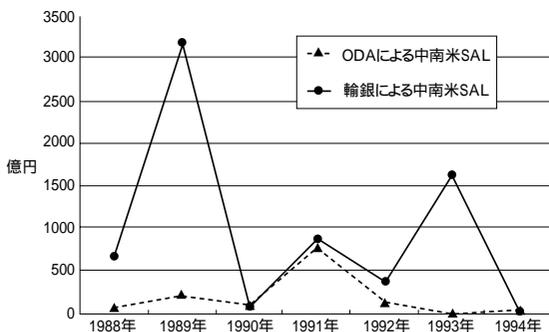
では、日本のODAとしては、構造調整借款という援助モダリティを選択することの妥当性をどう判断していたのだろうか。

これまでの調査から明らかに言えることは、まず、資金環流目的というミッションがあり、それを実行するにあたり、国際援助潮流を鑑み、世界銀行との協調融資のかたちで構造調整による融資案件が選択されていったという点である¹⁷。

だが、構造調整借款を実施するにあたり、日本国内でいくつかの疑問点が投げられていた。海外経済協力基金は1988年4月に経済部を設置している。その目的は構造調整借款に専門的に取り組むことであった。構造調整借款の専門知識や技術が不足する中で、世界銀行から専門家を採用するなどの工夫をしていた。当時、経済部長であった下村氏によれば、この時期は、構造調整借款を学ぶプロセスにあったという。プロジェクトを中心に援助を実施してきた職員たちにとって、構造調整借款のメカニズムや想定された効果について、必ずしも統べてを理解・納得していたわけではなく、またこれまでのプロジェクト援助と異なり相手国の政策課題により深く関与するため、構造的な問題解決の必要性を認めながらも違和感もあったという。

つまり、ODAの構造調整借款に着手した当初は学習過程にあり、ODAとして構造調整借款を用いることの妥当性や目的を定めかねていたのではないだろうか。1992年には、海外経済協力基金から世界銀行の構造調整借款に対する反論ペーパーが出されているが、ここで日本のODAとしての構造調整借款に対する意見が明示のかたちで出されている。その後、構造調整借款の実績が下降傾向になっていることを鑑みれば、必ずしも日本のODAとしては構造調整借款を手放しで受け入

図9 輸銀ODAによる中南米構造調整借款推移



(出所) 国際協力銀行HP「資金環流・資金協力計画」(1987-1998)、日本輸出入銀行海外投資研究所「海外投資研究所報」(1986-1995)、外務省「調整融資のレビュー報告書」(2005)より作成

れているわけではない。その一方で、資金環流措置を背景に当時、構造調整借款の実績額は急増しており、当時のODA総額に占める割合も13%以上と決して少なくない。したがって、日本のODAとして構造調整借款を選択したことの妥当性を検証しておく必要がないだろうか。

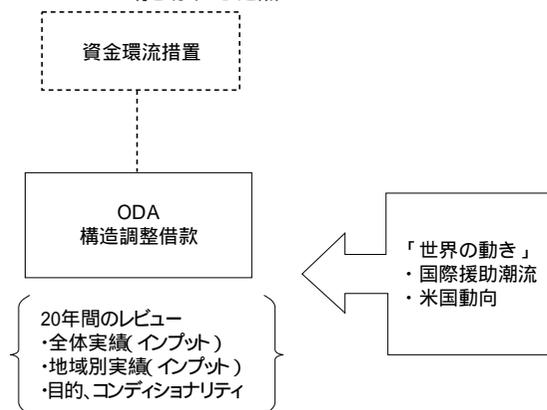
5. 政策評価の新たな課題

本論は、日本の構造調整借款のレビュー結果から得られた発見事項を更に発展させ、構造調整借款の背景にある日本政府の政策判断を探ろうとした。そこには日米貿易摩擦など国家の意思に直結する政策的背景がみえてきた。そこで、外務省の構造調整レビューと本論の相違点について明らかにしながら、政策評価の新たな可能性について考察してみたい。

(1) 構造調整（ODA）レビューと本論の相違

外務省の構造調整レビューでは、構造調整借款の拠出額の年代別増減、地域別傾向、目的やコンディショナリティの変化や傾向について明らかにした。同時に、構造調整20年間の動向に関係したと思われる、日本の動き、世界の動きを記した。但し、この点についてはあくまでも動向であって、確固たる政策判断として記したものはなかった

図10 外務省「調整融資のレビュー報告書」で明らかにした点



(出所) 筆者作成

(図10)。

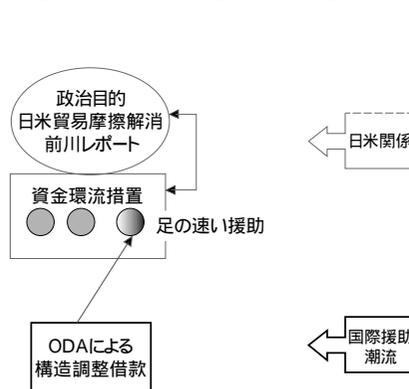
一方、本論では、ODAの構造調整借款は資金環流措置のパッケージのひとつとして選択されたという政策判断があったこと、さらにその上位の政策判断には日米貿易摩擦解消の手段として資金環流措置が作られたことが明らかにされている。したがって、当時の国際情勢、それを踏まえた日本の政治および政策目的との関係がより鮮明になった(図11)。

つまり、外務省レビューではODAの構造調整借款の動向と日本および関連の国際動向を暗に示唆するに留まっていたが、本論はこの点を明らかにしながら、ODAのより上位にある政策目的と判断を明らかにしたのである。では、このことは政策評価にとって何を示唆しているのだろうか。

(2) どのレベルの政策に着目するのか

図12は構造調整借款の上位にある政策目的、さらに上位にある政策目的との関係を図示したものである。ODAとしての構造調整借款の目的は、その上位にある資金環流措置の政策目的にとっては手段に位置付けられる。さらに、資金環流措置は、日米貿易摩擦解消という政策目的にとっては手段になっている。つまり、ODAとしての政策目的の上に更に上位の政策目的が存在している。我が国のODA評価ガイドラインに基づけば、まず当該案件とODA政策目的との整合性をもって

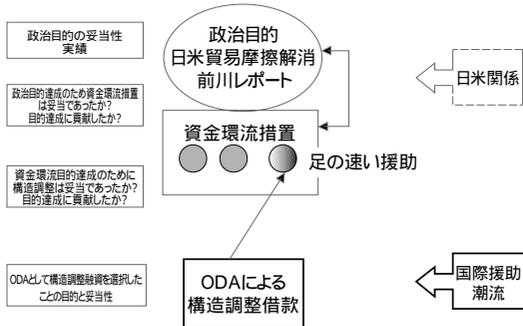
図11 本研究で明らかになったもの



(出所) 筆者作成

図12 今後の政策評価課題

外交手段、政治目的との関係が問われる



(出所) 筆者作成

その妥当性を判断することが求められる。

しかし、構造調整借款のようにその上に上位の政策が顕著に存在している場合、どのレベルの政策目的をもって政策評価に着手すべきなのだろうか。ODAはODAとしてその範疇で妥当性を判断していればよいという意見もあるが、このような国際情勢を見据えた大きな政治的判断の流れの中であって、ODAの政策目的に留まっては余りにも大きな政策視点を見逃してしまう。

だが、このような政策の上流部分に関する議論は政治学や政策分析の領域であるとして、評価という学問領域ではあまり着手してこなかったところではないだろうか。ささやかな試みであるが本論はODA目的とさらにその上流にある政策目的と串刺しにしてみることを試みた。

では、このような試みを可能にした分析視点を改めて顧みたい。

第1に時間軸で、20年間という長期の枠組みで構造調整借款供与額の増減の変化をみたことによる。増減の変化から、背後に何らかの政策判断があったのではないかと推測が成り立った。

第2に国際社会や日本政治・経済の動向にも着目したことである。本論は、特に日本が構造調整借款に着手する前後の政策動向に着目したが、国際社会の動向、すなわち債務危機とそれに対応した世銀などの国際機関のアクションとの間に数年の時間差があったためである。

第3にODA以外の国際協力スキームにも着目し

たことである。ODAの構造調整借款は、日本輸出入銀行による構造調整借款と国際機関向けジャパン・ファンドとの組合せによる資金環流措置のパッケージのひとつであった。

このことはODAの構造調整借款のみに着目していたのでは明らかにならなかった。同時にこのことは、行政府機関の縦割りの仕組みの中での評価やレビューでは容易に明らかでない点であることも示唆している。

これらの分析視点をあわせもつことによってODAのより上位にある政策の存在に届くことができた。1980年代当時の政策決定プロセスに伴う経験は、日米関係、国際機関との関わり方など豊富な教訓を含んでいる。このことは今後、我が国の外交政策、ODA政策を構築してゆく上で貴重な情報を提供しているのではないだろうか。

6. むすび

構造調整借款の背景にみえてきたのは、前川レポート、そして資金環流措置であった。前川レポートにより積極的な内需拡大策がとられてゆくが、これが後の経済バブルに発展したことは周知のことであるが、日本の援助にとっても大きな変化をもたらす出来事であった。このようにODA政策の上流に目を向けようとすると、省庁間の枠組みを超え政治的な判断まで見てゆくことになった。

構造調整借款のように、ODA政策を越えた政治や政策判断に直結する案件にとって、その妥当性を判断するには、ODA大綱や中期政策との整合性のみでは説明しきれない。また縦割りの仕組みの中での評価体制の限界を暗示しているのかもしれない。

ODAにとって教訓は何か。資金環流措置に押されるかたちで「学びながら実行する」状態であったため、構造調整借款開始当初は日本のODAとしての確たる目標を定めきれていなかったことである。その後、90年代に入り、日本は世界銀行の構造調整借款のあり方について反論ペーパーを出している¹⁸。日本のODAとしての考え方を明示

のかたちで出したものだ。また、本論では明らかにしていないが、何故、ODAの構造調整借款をアフリカを中心に始めたのか、日本政府の目的は何であったのかを明らかにすることは更なる課題である。

最近、アフリカを中心に国際援助界で一般財政支援という政策対話をベースにした援助形態が潮流になりつつあるが日本政府はこの新しい援助モデルの取り組みを検討している。そこでは援助の潮流との調和も重要であるが、同時に日本のODAとしての目的を明確に打ち出すことが必要であろう。だからこそ、構造調整借款という、当時、新モデルを用いた日本の取り組みの妥当性と効果について総括する必要があるのではないだろうか。

謝辞

本調査は20年前の政策と執行に関わられた方々の協力なしには実現できなかった。内海孚氏、下村恭民氏、幸島祥夫氏、前川鉄馬氏、宮元巖氏、広瀬哲樹氏はじめ御協力いただいた多くの方々改めて感謝致します。

注記

- 1 Structural Adjustment Loans (SALs), Sector Adjustment Loans (SECALs), Programmatic Structural Adjustment Loans (PSALs), Special Structural Adjustment Loans (SSALs), Subnational Adjustment Loans (SNALs)
- 2 外務省(2004)「調整融資レビュー」p.15、16の図表7,8,9からも、SALとSECALが年代を通じて、ほぼ同比率で実施されていることがわかる。
- 3 外務省「ODA50年の歩み」に掲載されている「日本のODA実績の推移(1965-2002)」「中期目標(第1次~第5次)」から算定した。
- 4 外務省構造調整借款レビュー報告書では世界銀行の構造調整融資と日本のそれとの傾向を比べているが、目的やコンディショナリティは同様の傾向を示すが、地域別にみると必ずしも世界銀行と一致しない。特にそれが顕著なのは東南アジア向け構造調整借款で日本は91年から97年まで構造調整借款を行っ

ていないが、世界銀行は実施しているというコンラストを示している。

- 5 広瀬氏、下村氏からのヒアリング。
- 6 内海孚氏 日本格付研究所取締役社長(1981年から86年まで大蔵省国際金融局長)、中平幸典氏 信金中央金庫理事長(92年から大蔵省国際金融局長)、幸島祥夫氏(株)T&Kパートナーズ代表取締役、元東京三菱銀行常務取締役、広瀬哲樹氏 元外務省審議官、下村恭民氏 法政大学教授、元海外経済協力基金職員、藤川鉄馬氏 住友信託銀行顧問、宮本巖(株)キューラメディクス 代表取締役会長
- 7 最重貧国の債務状況に着目すると、公的部門の債務残高は、1970年、53億ドルであるのに対し、民間部門では2億ドル、1980年は公的部門が4006億ドル、民間部門は43億ドルである(World Bank,1999)。従って、民間銀行からの低所得国への貸付は中所得国へのそれに比較して圧倒的に少ないといえる。
- 8 民間資金流入総額：1970~1974年は4,617百万ドルであったものが、1975~1979年には15,646百万ドルに増えている。出典：丸谷吉男(1989)「ラテンアメリカの債務危機と外国投資」アジア経済研究所、第21表 外部資金純流入と構成
- 9 外務省報告書では、第2次石油危機を債務危機発生の直接的原因として記したが、正確にはFRBのインフレ抑制政策により、実質金利が上昇したことで、民間銀行がラテンアメリカの債務国に対して新規融資を渋り始めたことにあると言ったほうが適当であろう。
- 10 債務国のひとつが、IMFに救済を求めると、IMFは構造調整借款プログラムを債務国と協議の上作成する。次に債権を有する銀行から構成されるBank Advisory Committeeのメンバーが当該政府と交渉し、資金ニーズの分析、リスケジュール、新規融資の交渉を取りまとめる。BACのミッションは、銀行の協調による金融支援により、債務国の経済の安定成長を促進し、国際金融の秩序と円滑化をはかるといったものであった(幸島2002、p.132)
- 11 BACの元、民間銀行は1983-1984年には53億ドルの新規融資を増額させたが、85年には8億ドル、86年には28億ドル減少と徐々に要請に応じなくなっていた。
- 12 ラテンアメリカ向け貸出しの内訳は、米国銀行が37%、日本銀行15%、英国14%、フランス10%であ

る。他国の銀行が切り詰める中で、邦銀だけが貸出しを増加させ、34年には290億ドルであったのが、86年には370億ドルとなった。

13 昭和61年 法人税法第52条、施例96.97

14 内海氏とのインタビューより。大蔵省広報誌「ファイナンス」1989年4月号p.5

15 内海氏、藤川氏とのインタビュー。

16 構造調整借款が始められたのは1986年、資金環流措置は1987年からである。当時の関係者によれば、86年当時、アフリカで構造調整借款を始めた頃は、援助潮流に合わせながら、新援助モダリティを学習するという目的もあった。その後、1988年にはアフリカ向け構造調整借款額が急増している。このことから、資金環流措置と構造調整借款の関係をみてとれる

17 構造調整借款に着手するにあたり内政干渉ではないかという議論もあったが、債務危機のような問題の抜本解決には政策対話も必要という意見もあった。そこで世界銀行と協調融資というかたちをとることによって、内政干渉への批判をかわし、政策対話の方法を学ぼうとする議論があった(広瀬氏からのヒアリングより。)

18 海外経済協力基金(1992)「世界銀行の構造調整アプローチの問題について」『基金調査季報』No.73

参考文献

藤川鉄馬(1989)「累積債務問題の新たな展開」『大蔵省ファイナンス』1989年4月号

海外経済協力基金(1992)『海外経済協力基金三十年史』

海外経済協力基金(1992)「特集1 世界銀行の構造調整アプローチの問題について」『基金調査季報』No.73

外務省(2005)『調整融資のレビュー 構造調整借款及びセクター調整借款の概観 報告書』

外務省『我が国の政府開発援助(上、下)』『政府開発援助白書』各年版

幸島祥夫(2002)『バランスシートによる日本経済分析』シーエーピー出版

国際協力銀行(2003)『海外経済協力基金史』

下村恭民、西垣昭(1993)『開発援助の経済学』有斐閣
下村恭民、中川淳司、齋藤淳(1999)『ODA大綱の政治経済学』有斐閣

世界平和研究所(1993)『中曽根内閣史』1巻~5巻

日本輸出入銀行『海外投資研究所法』(1986-1995)

平田喜彦(1993)『現在国際金融の構図』法政大学出版局

ペドロ=パブロ・クチンスキー著、渡辺敏訳(1990)『中南米債務 危機のメカニズムと打開策』サイマル出版会

ポール・ボルガー、行天豊雄著江澤雄一監訳(1992)『富の興亡 円とドルの歴史』東洋経済新報社

堀内昭義編(1991)『累積債務と財政金融』アジア経済研究所

丸谷吉男編(1989)『ラテンアメリカの経済危機と外国投資』アジア経済研究所

御厨貴著(2002)『オーラル・ヒストリー』中公新書

毛利良一(1995)「タンザニアにおける構造調整と社会開発」『日本福祉大学経済論集 第12号』

毛利良一(1988)『国際債務危機の経済学』東洋経済新報社

毛利良一(2001)『グローバル化とIMF・世界銀行』大月書店

SPA-6(2002). *Implementing the New Partnership.*

World Bank(1990). *Adjustment Lending Policies for Sustainable Growth, Policy Research And External Affairs The World Bank*(September).

World Bank(1994). *Adjustment in Africa.* The World Bank and Oxford University Press.

World Bank(2001). *Adjustment Lending Retrospective Final Report.* Operations Policy and Country Services (June 15).

World Bank(2003). *Programmatic Adjustment Lending Retrospective,* Operations Policy and Country Services (July 11).

World Bank(2004). *Strategic Framework for Assistance to Africa.* the World Bank.

(2006.1.30受理)

The Twentieth-Year Review of Japanese Structural Adjustment Loans: Political Background and Decision-Making of the Japanese Government

Yayoi Tanaka

University of Tokyo

yayoi-tanaka@civil.t.u-tokyo.ac.jp

Keywords

Structural Adjustment Loans, US-Japan trade imbalance,
Maekawa Report, Financial Recycling Plan, Policy Evaluation

The Japanese Ministry of Foreign Affairs recently conducted the first comprehensive review of Structural Adjustment Loans (SAL) provided by the Japanese government. For the roughly two decades from 1986 to the present, approximately 9,676 billion yen has been disbursed as SAL. This paper attempts to uncover the political background for the decision to start SAL in the 1980s. The author focuses on the fact that the Japanese government faced a serious problem from the US-Japan trade imbalance at the time. Under Prime Minister Yasuhiro Nakasone's initiative, it announced the Financial Recycling Plan with the aim of circulating about US\$65 billion in Japanese money to international institutions such as the IMF and World Bank, and to the developing countries, which were facing serious debt problems. SAL was selected as one of the components of the Recycling Plan, as it allowed the quicker disbursement of loans than project loans. The question here is the rationale for the selection of SAL as the modality for ODA. However, it is impossible to evaluate the relevance of the SAL program, which had a political mission beyond the purpose of simply providing ODA, under the current guideline for evaluating ODA. This paper raises the question about how to evaluate the program that is directly connected to the grave political purposes.

【研究論文】

標準化した効果単位を用いる定量的評価手法の提案 - 留学生政策と地域展開型プログラムへの適用例

佐藤 由利子

東京工業大学

yusato@ryu.titech.ac.jp

要 約

本研究では、非介入人口と比べた介入人口の指標（群）の変化率から標準化した効率を算出し、effectという単位を与え、この標準化した効果に介入人口を乗じてインパクトを、また、インパクトを投入金額で除して効率を測定する手法を提案する。更に、この手法を、留学生政策と地域展開型プログラム協力の評価に適用した事例を紹介する。

本手法の主な利点としては、単位換算が難しい介入の効果を、effectという標準化した単位で定量的に表現することにより、インパクト、効率についても定量的に測定することが可能になった点が挙げられる。このため、同じ指標（群）を使って評価する場合には、標準化した効果、インパクト、効率について、異なる介入主体間の比較や、異なる対象に対する介入の比較、また、経年変化を調べる事が可能となる。評価結果の差異は、主に、介入する側と介入される側の条件の違いによって生じるため、評価結果を比較することにより、効果的、効率的な介入のあり方について、また、介入対象人口における促進要因、阻害要因について、考察する糸口が与えられる。

キーワード

標準化効果、定量的評価、効率、インパクト、結果重視の評価

1. はじめに

政策、プログラム、プロジェクトなど、対象人口の経済、生活、能力、環境等の向上を目指す介入の効果、インパクト、効率を定量的に評価するには、どのような標準化された方法があるだろうか？

筆者は政策評価マトリックス (Policy Evaluation Matrix, PEM) という枠組みを用いて、インドネシアとタイの日本留学生、米国留学生、非留学生

のアンケートの比較分析により、日本の留学生政策について、人材養成（開発）と日本との友好促進（国益）という2つの視点からの評価を行ってきた。

本稿では、この手法を改良し、留学生政策、地域展開型プログラムなどの介入において、非介入人口と比べた介入人口の指標の変化率から一人当たりの標準化した効果を算出し、これに介入人口を乗じてインパクトを、また、インパクトを投入金額で除して、効率を測定する手法を提案する。

2. 過去の開発評価手法のレビューと本研究のアプローチ

過去の開発援助、開発介入に関する評価は、主に開発の視点から行われてきており、特に経済協力開発機構（OECD）加盟国による開発援助については、開発援助委員会（DAC）が開発援助評価原則を定め（OECD 1991）、これに則って各国のODA評価が行われてきた。

評価手法としては、ドイツのGTZで開発されたZOPP（*Ziel Orientierte Projekt Planung* or Object-Oriented Project Planning）手法に基づく、ロジカルフレームワークを用いたプロジェクトの計画、モニタリング及び評価が、OECD加盟国及び国際機関によって広く採用されている。日本においては、国際開発高等教育機構（FASID）が中心となり、ZOPP手法を改良したPCM（Project Cycle Management）手法を開発し（1996）この手法を用いた開発プロジェクト評価が、開発援助機関で広く用いられている。PCM手法では、PDM（Project Design Matrix）というロジカルフレームワークを用いるが、三好（2002、2003）は、このPDMの左列を集積したPTM（Program Theory Matrix）によって、複数のプロジェクトから構成されるプログラムのメタ評価を行うことを提案している。

筆者はPDMを上位の政策レベルにシフトした政策評価マトリックス（PEM）という枠組みを用いて、留学生政策等の開発援助を、開発と国益の2つの観点から評価することを提案してきた（佐藤 2002a、2002b、2003、2004、2005、Sato 2005）。その基本となる考え方や利点は、次の通りである。

- (1) PDMを上位の政策レベルにシフトして作成したPEMは、PDMと基本的に同じ論理構造をもち、DAC開発援助評価原則に示す目標達成度、妥当性、インパクト、効率性、自立発展性の5つの観点からの総合的な評価を行うことが可能となる。
- (2) PEMは政策に影響を与える、あるいは政策から派生するほぼすべての要素を列挙しており、評価者と政策策定者、その他の関係者が、政策のセオリーを確認し、議論するプラットフォームとして有効である。

(3) PEMに示した指標及び指標データ入手手段に沿って調査項目を組み立てることにより、評価者は、評価調査を計画することが容易となる。

(4) 日本の留学生政策は1954年以降、留学生送出国の人材養成（開発）と親日家養成（国益）の2つを政策目標として、政府開発援助（ODA）の一環として実施されてきたが、他のODA政策や事業においても、開発と国益の2つの観点から、同様の形での評価が可能である。

近年、世界銀行等を中心に、主な関連指標の変化に着目し、開発のプロセスより、結果を重視するモニタリング及び評価手法が広がりつつある（CIDA 2000, Kusek & Rist 2004）。この手法の利点としては、指標測定により、評価結果が比較的簡単に入手できることが挙げられる。他方、評価結果の背景にある因果関係は、指標だけでは解明することはできない。このため、結果重視の評価手法においても、PDMやPEMのようなロジカルフレームワークを用いた分析を併用することが必要となる。

筆者はPEMを用いて、インドネシアとタイに対する日本の留学生政策の評価を、主に世界最大の留学生受入国である米国との比較において行ってきた。このように、他国と比較して日本の政策の相対的達成度を評価する手法のメリットとしては、日本の政策の特色、強みと弱みが明確になる点が挙げられる。他方、デメリットとしては、政策の背景も中身も違うものを比較しているため、効果やインパクトを、必ずしも正確に測定していない点が挙げられる。この観点からは、日本留学生者と非留学生者（留学せずに自国の大学・大学院を卒業した者）の回答値や留学しなかった場合とを比較する方法が、有効と考えられる。

これらの点に鑑み、本稿では、留学生者と非留学生者の平均回答値、あるいは留学しなかった場合と比較することにより、インドネシアとタイに対する、日本の留学生政策の効果、インパクト、効率性を再測定する。具体的には、留学しなかった場合、あるいは非留学生者と比較した日本留学生者の平均回答値から標準化効果を、標準化効果に留学生数を乗じることによりインパクトを、更に、インパクト

トを投入金額で除すことにより、効率を算出する。

また、この手法の根拠となる基本的考え方を整理し、非介入人口と比較した介入人口における主要な指標の平均変化率と、介入人口の規模、投入金額に着目して、介入の効果、インパクト、効率を定量的に計測する手法として一般化する。更にこの手法を、保健分野の地域展開型プログラムの評価に適用した事例を紹介する。

本稿で提案する手法は、結果重視の評価手法の流れを汲み、指標変化に着目するものであるが、PEM等のロジカルフレームワークを併用することにより、投入、外部条件など、介入に影響を与える要素を明示し、成果、目標、上位目標の関係を整理できるため、当該指標変化をもたらした因果関係の考察も比較的容易である。また、介入や評価の枠組みを、介入主体が、介入人口や関係者と共有することが可能となる。

3. 効果及び効率測定のための基本的考え方と留学生政策評価への適用

本節では、非介入人口と比較した介入人口における指標の変化率から、標準化した効果を算出し、更に、人口規模、投入金額を勘案して、インパクトと効率を定量的に評価する手法の基本的考え方を紹介する。更に、この手法を、日本の留学生政策評価に適用した事例を紹介する。

本稿で提案する評価手法の前提となっている基本的考え方は、次の通りである。

ある介入において、その目標の達成を端的に示す指標（群）を設定した場合、非介入人口と比較した、介入対象人口における当該指標（群）の平均変化率は、介入がもたらす効果を示しており、次の数式で表すことができると考えられる。

介入の効果を、非介入人口における指標平均値M'と比較した介入人口における指標平均値Mの変化率と捉えると、
 <ストック指標においては>

$$\text{効果} = (M/M') - 1 = (M - M') / M'$$

 <フロー指標においては>
 介入前のベースラインの指標値をBとすると

$$\text{効果} = (M - B) / (M' - B) - 1 = (M - M') / (M' - B)$$

 ただし、M'=Bの場合には、

$$\text{効果} = (M - B) / M' = (M - M') / M'$$

経済学用語から、ストック指標は中長期的な特性を表す指標、フロー指標は1年程度の短期間に变化する指標と定義する。小野（2004）は、フロー指標とストック指標に関連し、2つの目標達成度の算出式を挙げている。

目標達成度 = 実績値/目標値

目標達成度 = (実績値 - 基準値) / (目標値 - 基準値)

筆者は、この小野の考え方を応用し、ストック指標については、非介入人口と介入人口における指標平均値を比較することにより、また、フロー指標については、非介入人口と介入人口の指標が、介入前のベースラインの指標値から変化し度合いの比較から、それぞれ介入の効果を出せることと考えた。

介入目標の達成を示す複数の指標がある場合には、まず、ストック指標とフロー指標のどちらを使用するかを選択し、次に、選択した指標について、指標間の重み付けが判明している場合には、それら指標の変化率の加重平均から、また、指標の重み付けが不明な場合には、便宜的に、それら指標の変化率の平均から、効果を算出することとする。

上記数式から算出される効果を示す値は、当該介入により、介入人口一人当たりどの程度の変化が起きたかという「一人平均の変化の度合い」を示す数値である。介入の種類によって、効果を金額などの単位付で表す場合もあるが（例えば一人当たり1万円の経済効果といった表現）、多くの介入においては、効果を単位換算することは容易ではない。

このため筆者は、上記数式から算出される効果を示すeffectという単位を新たに設定し、介入がもたらす変化として、標準化して表す。例えば、非介入人口と比べ、介入人口において平均1%の指標変化があった場合には0.01effect、という形で表記する。

ベースラインで平均100万円の売り上げが、介入人口において101万円となり、非介入人口では100万円のままであった場合、効果は次のように算出される。

$$(M - B) / M' = (M - M') / M' = (101\text{万円} - 100\text{万円}) / 100\text{万円}$$

$$= 0.01 \text{ effect}$$

先に示した一人平均1万円の経済効果という表現は、上記数式の分子の部分を表している。このことから、筆者の提案するeffectという単位は、分母のM'（ストック指標の場合とフロー指標でM'=Bの場合）あるいはM'-B（フロー指標においてM'≠Bでない場合）を1に置き換えた場合の、標準化した効果を表す単位であることがわかる。

効果について標準化した値を使う利点としては、複数の指標を使用するなど、単位換算が難しい介入の効果を表現できることが挙げられる。同じ指標（群）を使う場合、異なる介入主体による効果の比較や、異なる対象に対する効果の比較、また、効果の経年変化を調べることが可能となる。

次に、標準化効果単位に基づきインパクトを測定する方法を示す。インパクトは、介入が、介入人口に対してもたらした変化の総量、つまり、非介入人口と比較した対象人口個々の指標変化（＝個別効果）の総和であると考え、非介入人口と比べた介入人口の指標平均の変化率（＝標準化効果）に介入人口を乗じた次の数式で表すことができると考えられる。インパクトの単位は、effect・人となる。

介入人口個々の指標値を X_1, X_2, \dots, X_n 、各指標値の該当人数を f_1, f_2, \dots, f_n 、対象人口を f 、介入人口における指標平均値を M 、非介入人口における指標平均値を M' とすると、インパクトは

<ストック指標においては>

$$\frac{f(X_1 \cdot M')}{M'} + \frac{f(X_2 \cdot M')}{M'} + \dots + \frac{f(X_n \cdot M')}{M'} = \sum_{i=1}^n \frac{f(X_i \cdot M')}{M'} = \frac{f(M \cdot M')}{M'}$$

=効果 × 介入対象人口

<フロー指標においては>

$$\frac{f(X_1 \cdot M')}{(M' - B)} + \dots + \frac{f(X_n \cdot M')}{(M' - B)} = \sum_{i=1}^n \frac{f(X_i \cdot M')}{(M' - B)} = \frac{f(M \cdot M')}{(M' - B)}$$

=効果 × 介入対象人口

効率については、介入人口における変化の総量であるインパクトを、すべての投入を金額換算した投入金額で除することにより、次の数式で求めることができると考えられる。効率は、介入人口一人当たりの投入金額に対する指標の変化率を示すとも捉えられる。単位は、effect・人/金額単位として表される。

効率 = 効果 × 対象人口 / 投入
 = 効果 / (投入 / 対象人口)
 = 効果 / 一人当たり投入
 = 指標の変化率 / 一人当たり投入
 投入金額を I で表すと、効率は、
 <ストック指標においては>

$$\frac{f(M \cdot M')}{M' I} = \frac{(M \cdot M') \times f}{M' I}$$

<フロー指標においては>

$$\frac{f(M \cdot M')}{(M' - B) I} = \frac{(M \cdot M') \times f}{(M' - B) I}$$

以上のような考え方に基づき、インドネシアとタイに対する日本の留学生政策の効果、インパクト、効率を再測定することとしたい。

日本の留学生政策は、1954年に国費留学生制度が開始して以降、留学生送出国の人材養成と日本との友好促進という2つの政策目標の下に実施されており（佐藤2004:54-55）、これらの政策目標に対応したPEM（1954年～2001年を想定）は、表1、表2のとおりである。

人材養成と友好促進に関する留学生政策の効果と非留学生（または留学しなかった場合）の指標と比較した次の数式から、標準化した効果を算出する。留学に関する指標は、留学により中長期的に生じた変化を示しているため、ストック指標と考えられる。

留学生政策の標準化効果 = (日本留学者の指標平均値 / 非留学者または留学しなかった場合の指標平均値) - 1

また、留学生政策のインパクトは、平均留学期間を4年と想定すると¹⁾、来日から4年を経過した留学修了時点から現れると考えられる。このため留学生政策が本格化した1954年以降、筆者が調査

表1 留学生政策に関する政策評価マトリックス(PEM):人材養成の視点から

要約 (Summary)	指標	指標データ入手手段	外部条件
(上位政策目標)留学生送出し国の経済・社会が発展する	経済成長率、一人当りGDP等	統計資料	
(政策目標)日本の留学生政策により留学生送出し国の発展に貢献する人材が養成される	留学しなかった場合と比較した帰国留学生の職位と収入の向上度合い	1. アンケート調査 2. 帰国留学生職場での聞き取り調査	
(政策成果) 1. 優秀で意欲の高い者が留学する 2. 留学中に十分な知識と技能を習得する 3. 帰国留学生が留学で習得したことを仕事で活用し、また他の人々に伝達する	1. 留学動機、留学中の勉学態度、日本語能力 2. 留学教育に対する満足度、学位取得状況 3. 職場での留学成果の活用状況、知識・技能の伝達状況	1. アンケート調査、日本大使館での調査 2. アンケート調査、面接調査 3. アンケート調査、職場でのヒアリング	1. 留学生の多くが帰国する 2. 帰国留学生に雇用機会が提供される
(政策アウトプット) 留学生の日本への受入れ	留学生のべ受入れ人数(人年)	統計資料	
(主な施策) 1-1 日本留学情報の提供 1-2 国費奨学生募集・選考・受入れ 2-1 留学生に対する教育・研究指導 2-2 私費留学生への支援 2-3 留学生宿舍の確保、医療費補助 3-1 元留学生へのフォローアップ	(主な施策担当機関・実施機関) 1-1外務省(在外公館他)、文部科学省(AIEJ) 1-2外務省(在外公館)、文部科学省 2-1文部科学省(高等教育機関) 2-2文部科学省(AIEJ) 2-3文部科学省(AIEJ、内外学生センター) 3-1外務省(在外公館他)、文部科学省(AIEJ)		

(注1)AIEJは文部科学省管轄の財団法人日本国際教育協会(現・日本学生支援機構JASSO)を指す。

(注2)本PEMIは国費及び私費(自費・民間奨学金)留学生を念頭に作成したため、外国政府派遣留学生やJICA長期研修員には適用しない。

(出所)筆者作成

表2 留学生政策に関する政策評価マトリックス(PEM):友好促進の視点から

要約	指標	指標データ入手手段	外部条件
(上位政策目標) 日本との友好関係が促進される	対日感情、交流実績	統計資料、日本大使館での調査	
(政策目標)日本の留学生政策により、元留学生が親日家となる	非留学者と比較した帰国留学生の親日度の変化	アンケート調査	
(政策成果) 1. 優秀で意欲の高い者が留学する 2. 留学中の教育・生活に満足する 3. 留学中に良好な師弟・友人関係を築き、留学後も継続する 4. 元留学生が日本との友好促進活動に参加する	1. 留学動機、留学中の勉学態度、日本語能力 2. 留学中の教育・生活に対する満足度 3. 師弟・友人関係の構築と継続 4. 日本との友好促進活動への参加率	1. アンケート調査、日本大使館での調査 2.3.4. アンケート調査	日本との間に大きな外交的障害がない
(政策アウトプット) 留学生の日本への受入れ	留学生のべ受入れ人数(人年)	統計資料	
(主な政策) 1-1日本留学情報の提供 1-2国費留学生募集・選考・受入れ 2-1留学生に対する教育・研究指導 2-2私費留学生への支援 2-3留学生宿舍の確保、医療費補助 3-1留学生との交流 4-1元留学生へのフォローアップ	(主な施策担当機関・実施機関) 1-1外務省(在外公館他)、文部科学省(AIEJ) 1-2外務省(在外公館)、文部科学省 2-1文部科学省(高等教育機関) 2-2文部科学省(AIEJ) 2-3文部科学省(AIEJ、内外学生センター) 3-1文部科学省(高等教育機関、留学生交流推進会議、国際交流団体) 4-1外務省(在外公館他)、文部科学省(AIEJ)		

(注1)、(注2)及び(出所)は表1と同じ

を行った2001年時点までに現れた留学生政策のインパクトは、標準化効果に1958年～2001年の留学修了者人数を乗じた次の数式で算出できると考えられる。

$$1954\text{年}-2001\text{年の留学生政策のインパクト} = \text{留学生政策の標準化した効果} \times 1958\text{年}-2001\text{年の留学修了者人数}$$

実際は、1958年～2001年の留学修了者数の正確な統計が入手困難なことから、1954年～1997年に来日した留学生全員が留学を修了し、2001年時点で生存していると仮定し、1954年～1997年の留学生受入れ人数で代替して、インパクトを算出することとする。

$$1954\text{年}-2001\text{年の留学生政策のインパクト} = \text{留学生政策の標準化効果} \times 1954\text{年}-1997\text{年の留学生受入れ数}$$

また政策効率については、インパクトを、1954年～1997年の投入金額（各年の予算投入をデフレートした上で累計した額）で除することにより算出する。効率は、1954年～1997年の留学生一人当たりの投入金額に対する、指標の変化率を示すとも捉えられる。

$$1954\text{年}-2001\text{年の留学生政策の効率} = \text{留学生政策の標準化効果} \times 1954\text{年}-1997\text{年の留学生受入れ人数} / 1954\text{年}-1997\text{年の投入} = \text{標準化効果} / (1954\text{年}-1997\text{年の投入} / 1954\text{年}-1997\text{年の留学生受入れ人数})$$

上記の算出式を用いて、筆者が2001年、2002年にインドネシアとタイで実施したアンケートの結果から²、人材養成と友好促進の2つの目標に関して、留学生政策の標準化効果、インパクト、効率を計算する。

インドネシアとタイの日本留学生と非留学生の主な属性は以下のとおりである（佐藤2002a、2000b、2003、2004、Sato 2005）。

<インドネシア>

日本留学生：男女比（78：22）、生年平均1957年、職業（大学教員46%、公務員・準公務員36%、企業勤務者17%）
非留学生：男女比（78：22）、生年平均1961年、職業（大学教員40%、公務員・準公務員60%）

<タイ>

日本留学生：男女比（61：39）、生年平均1957年、職業（大学教員26%、公務員・準公務員21%、企業勤務者55%）
非留学生：男女比（49：51）、生年平均1967年、職業（大学教員44%、公務員・準公務員1%、企業勤務者51%）
職業は複数回答を含む。

非留学生の回答者は、日本留学生に近い職業、男女比、年齢構成となるよう配慮したが、インドネシアでは、日本留学生に比べ、非留学生の公務員・準公務員比率が高い。また、タイにおいては、日本留学生に比べ、非留学生の平均年齢が若く、公務員・準公務員比率が小さく、大学教員比率が高い。日本留学生と非留学生を比較する際には、これらの点に留意する必要がある。

また、国費留学生と私費留学生（＝自費・民間奨学金受給者）については、投入予算が異なるため、両者を区別して効果、インパクト、効率を測定した。

表1、2の下段に一部示すとおり、国費留学生施策としては、国費留学生の募集・選考、受入れが、私費留学生への施策としては、授業料減免、学習奨励費などの支援策が、また両者に共通する施策として、留学情報の提供、留学生に対する研究・教育指導、留学生宿舍の確保、医療費補助などが挙げられる。

人材養成に関する主な指標としては、職場における日本留学生と非留学生の業績の差、留学しなかった場合と比較した、地位や給与の上昇率などが考えられる。これらの指標のうち、筆者が実施したアンケート調査では、日本留学生に対して、留学しなかった場合と比較した地位や収入の変化率をたずね、次のような回答平均値が得られた³。

<留学しなかった場合と比較した地位の向上割合>
インドネシアの国費留学生46.4%、
私費留学生35.0%（有意差）
タイの国費留学生41.2%、
私費留学生45.8%（有意差なし）
<留学しなかった場合と比較した収入の向上割合>
インドネシアの国費留学生31.1%、
私費留学生28.8%（有意差なし）
タイの国費留学生26.2%、
私費留学生32.8%（有意差なし）

t 検定の結果、インドネシアでは、国費と私費留学生の間で、地位の向上には5%水準で有意差(以下、有意差はすべて5%水準で計算)があるが、収入の向上には有意差が見られなかった。この背景として、インドネシアでは、国費留学生制度が、インドネシア政府により、大学教員や公務員の人材養成の機会として活用され、復職後、留学経験を活かして組織で昇進する機会を与えられていること、しかし、給与は民間に比べ少なく、私費留学生(企業勤務者がやや多い)とさほど大きな差が生じないことが挙げられる。また、タイでは、国費と私費留学生の回答に有意差は見られず、数値自体は、私費留学生がやや高い。私費留学生は、国費に比べ、民間企業、特に日系企業に就職する者の割合が高く、日本留学経験をビジネスで活かして、国費留学生同様、地位や収入の向上を達成していると考えられる⁴。

地位や収入の向上は、業績評価以外の要素(例えば派閥等)によって左右されることもあり、また、「留学しなかった場合」との比較は、元留学生の認識に基づく数値であるため、指標として取り扱うには注意が必要だが、ここでは便宜上、留学しなかった場合と比べた地位と収入の向上率を、人材養成に関する政策効果を同等の重みで示す指標と見なし、両者の平均から、下記のように標準化した効果を算出した。

<人材養成に関する標準化政策効果>	
インドネシアの国費留学生	0.388 effect
インドネシアの私費留学生	0.319 effect
タイの国費留学生	0.337 effect
タイの私費留学生	0.393 effect

次に、人材養成の観点からの留学生政策のインパクトを算出する。1954年から1997年の間に、インドネシアからの国費留学生は1,742名、私費留学生は1,412名、また、タイからの国費留学生は2,396名、私費留学生は2,469名と推計される⁵。人材養成に関するインパクトは、国費と私費の標準化政策効果に、1954年～1997年の国費と私費の留学生数を乗じて、下記のように算出される。

<人材養成に関する政策インパクト>	
インドネシアの国費留学生	0.388effect × 1,742人 = 675.9 effect・人
インドネシアの私費留学生	0.319 effect × 1,412人 = 450.4 effect・人
タイの国費留学生	0.337 effect × 2,396人 = 807.5 effect・人
タイに対する私費留学生	0.393 effect × 2,469人 = 970.3 effect・人

政策効率は政策インパクトを、1954年～1997年の政策投入額で除して算出する。なお、私費留学生については、授業料等の収入を考慮するため、各年の予算投入から授業料等の収入を差し引いた「純投入」でインパクトを除いて効率を算出する⁶。純投入がマイナスの場合は、国全体として、収入が投入を上回っていることを示す。人材養成に関する政策効率は、次のように算出される。

<人材養成に関する効率>	
インドネシアに対する国費留学生	675.9 effect・人 / 2,145百万円 = 0.315 effect・人/百万円
インドネシアに対する私費留学生	450.4 effect・人 / (-698百万円) = -0.645 effect・人/百万円
タイに対する国費留学生	807.5 effect・人 / 3,178百万円 = 0.254 effect・人/百万円
タイに対する私費留学生	970.3 effect・人 / (-1,121百万円) = -0.866 effect・人/百万円

私費留学生については、1954年～1997年の授業料等収入が予算投入を上回り、純投入がマイナスであるため、正確な効率は計算できないが、プラスの政策インパクトと収入の両方をもたらしていることがわかる。特に、タイの私費留学生においては、もたらずインパクトも収入も大きい。また、国費留学生については、政策効果の違いから、タイよりインドネシアに対する政策効率がよいことがわかる。

友好促進に関する政策効果は、表2のPEMに記載した「元留学生が親日家となる」という政策目標に基づき、「日本人が好き」という設問に対する、非留学生と比べた日本留学生の回答変化率を指標として用いることとする。

「日本人が好き」という設問⁷では、1～5の5段階評価で下記のような回答平均値が得られた。t

検定によりインドネシアでもタイでも、日本留学生と非留学生の間に有意差が見られ、留学経験による親日度の向上が確認された。他方、国費と私費留学生の間には、有意差は見られなかった。

「日本人が好き」に対する5段階評価の回答平均値
 インドネシア：国費留学生4.17、私費留学生3.94、非留学生3.80
 タイ：国費留学生3.61、私費留学生3.74、非留学生3.41

次に、上記の5段階評価の回答値を原点を持つ0～4の5段階評価の回答値に直し（各回答値から1を減）さらに0～4の回答が各々0%、25%、50%、75%、100%という比例尺度に対応していると見なして、非留学生と比べた日本留学生の親日度変化率から、標準化効果を次のように算出した。

<友好促進に関する標準化政策効果>
 インドネシアの国費留学生
 $\{(4.17-1)-(3.80-1)\}/(3.80-1) = 0.132\text{effect}$
 インドネシアの私費留学生
 $\{(3.94-1)-(3.80-1)\}/(3.80-1) = 0.05\text{ effect}$
 タイの国費留学生
 $\{(3.61-1)-(3.41-1)\}/(3.41-1) = 0.083\text{effect}$
 タイの私費留学生
 $\{(3.74-1)-(3.41-1)\}/(3.41-1) = 0.137\text{effect}$

この標準化効果に基づき友好促進に関する政策インパクトと効率を計算すると、次のようになる。

<友好促進に関する政策インパクト>
 インドネシアの国費留学生
 $0.132\text{effect} \times 1,742\text{人} = 229.9\text{ effect} \cdot \text{人}$
 インドネシアに対する私費留学生
 $0.05\text{ effect} \times 1,412\text{人} = 70.6\text{effect} \cdot \text{人}$
 タイの国費留学生
 $0.083\text{effect} \times 2,396\text{人} = 198.9\text{effect} \cdot \text{人}$
 タイの私費留学生
 $0.137\text{ effect} \times 2,469\text{人} = 338.3\text{effect} \cdot \text{人}$

<友好促進に関する政策効率>
 インドネシアに対する国費留学生
 $229.9\text{ effect} \cdot \text{人} / 2,145\text{百万円} = 0.107\text{ effect} \cdot \text{人/百万円}$
 インドネシアに対する私費留学生
 $70.6\text{effect} \cdot \text{人} / (-698\text{百万円}) = -0.101\text{ effect} \cdot \text{人/百万円}$

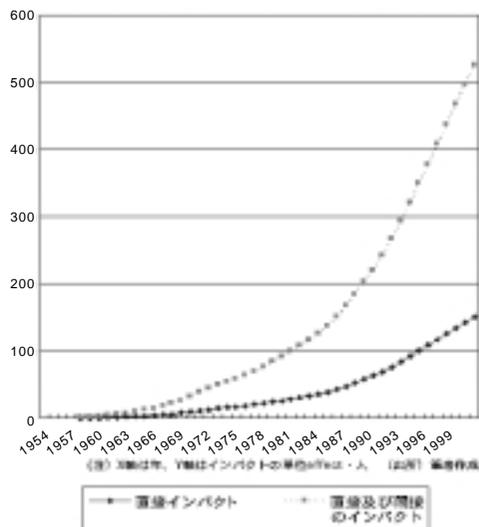
タイの国費留学生
 $198.9\text{effect} \cdot \text{人} / 3,178\text{百万円} = 0.063\text{ effect} \cdot \text{人/百万円}$
 タイの私費留学生
 $338.3\text{effect} \cdot \text{人} / (-1,121\text{百万円}) = -0.302\text{ effect} \cdot \text{人/百万円}$

以上より、私費留学生については、純投入がマイナスのため正確な効率は計算できないが、友好促進に関しても、プラスのインパクトと収入の両方をもたらしており、特にタイの私費留学生におけるインパクトと収入が大きいことがわかる。また、国費留学生については、標準化効果の違いより、タイよりインドネシアにおいて、政策の効率がよいことがわかる。

人材養成と友好促進に関しては、使用した指標が異なるため、効果、インパクト、効率を直接比較することはできない。しかし、同じ目標に関する介入を、同じ指標を使って評価する場合には、この手法を用いることにより、異なる介入主体間で、また、異なる介入対象間で、効果、インパクト、効率を比較することが可能となる。また、経年変化を調べることも可能となる。

図1は1954年～2001年の、タイの国費留学生における、友好促進に関するインパクトの変化を表したものである。

図1 親日家養成に関するインパクトの経年変化（タイに対する1954年-2001年の国費留学生政策）



指標の変化は、留学生受入れ開始後、平均留学期間4年の経過後に表れると想定されることから、例えばA年のタイの国費留学生における友好促進に関するインパクトは、次の数式で表すことができると考えられる。

$$\text{A年のタイに対する国費政策の友好促進に関する直接インパクト} = 0.083\text{effect} \times \{1954\text{年の留学生受入れ数} + 1955\text{年の留学生受入れ数} + \dots + \text{A-4年の留学生受入れ数}\} = 0.083\text{effect} \times (\text{1954年からA-4年までの受入れ人数累計})$$

図1の実線の折れ線は、タイの国費留学生における、友好促進に関する政策の直接インパクトの経年変化を示している。ただし、この直接インパクトは、留学修了者における指標変化の総和のみを計算しており、留学修了者が周辺の人にもたらず間接インパクトを考慮していない点に留意する必要がある。

菊田と牟田（2004）はアジア生産性機構による地域社会総合開発研修プログラムの研修参加者調査から、講演、セミナー、論文等の手段により研修成果を伝達した人数を、年間2.47人と計算している。このことから、日本留学修了者が、日本に関する好意的な情報を周囲に伝達した人数も、年間同程度であると仮定すると、友好促進に関する間接インパクトは次の数式で表すことができる。図1の破線の折れ線は、これら数式に基づいて算出した、タイの国費留学生における、友好促進に関する政策の直接と間接を合わせたインパクトの経年変化を示している。

$$\text{A年の国費留学生政策の親日家養成に関する間接インパクト} = 0.083\text{effect} \times (\text{1954年からA-4年までの留学生受入れ数}) \times 2.47\text{人}$$

以上、留学生政策の人材養成と友好促進に関する効果、効率、インパクトについて、留学しなかった場合、あるいは非留学者と比べた日本留学生の回答の変化率から、定量的に測定した事例を紹介した。

留学生政策は、日本のODAの一環として実施されているが、同じくODAとして実施されている研修員受入れ事業についても、同様の手法での定量的評価が可能であると考えられる。

ODAによる研修員受入れ事業を、国内最大規

模で実施している国際協力機構（JICA）は、ホームページに次のように記載している。

「研修員受入れ事業は、開発途上国の人材育成を目的として、開発途上国の今後の中核的な行政官、技術者、研究者などを『研修員』として招き、その国が必要とされている知識や技術を伝えるもので最も基本的な『人づくり』事業の一つとして位置づけられている。」

人材養成が事業目標であると明言されているが、研修員が日本を理解し、親日的になることを通じた開発援助の円滑化や日本との友好促進も、明文化されていないものの、援助国側の達成したい目標の1つと推定することができる。このため、留学生政策評価で使用した人材養成と親日家養成の2つの視点からの評価手法を適用することは可能であると考えられる。

研修員受入れ事業については、公開された評価資料が少ないため、本手法の適用事例を示すことはできないが、将来、留学生政策との比較を含めて、実証的な評価を行うことは、十分意義があることと思われる。

4. 地域展開型案件における応用：リプロダクティブ・ヘルス・プログラムの事例

本節では、前節で紹介した介入に関する効果、インパクト、効率の測定手法を、地域に展開するリプロダクティブ・ヘルス（RH）プログラムに応用した事例を紹介する。

地域展開型プログラム/プロジェクトにおける介入の効果については、非介入地域（介入地域とよく似た条件の地域）と比べた介入対象地域の指標の平均変化率から求めることができると考えられる。

例えば、農業分野の協力では農業生産性に関する指標、保健分野の協力では保健指標、教育協力では教育指標の、非介入地域と比較した介入地域における変化率が、開発効果を示すと考えられる。また援助国の国益の視点からは、非介入地域と比較した介入地域における、援助国への好感度に関する指標変化が、介入効果を示すと考えられる。

インパクトは標準化効果に介入対象人口を乗じ

ることにより、効率はインパクトを投入金額で除すことにより、次の数式で求めることができると考えられる。

$$\begin{aligned} \text{インパクト} &= (\text{非介入地域と比較した}) \text{介入地域の指標} \\ &(\text{群}) \text{の平均変化率} \times \text{対象人口} = \text{標準化効果} \times \text{対象人口} \\ \text{効率} &= \text{標準化効果} \times \text{対象人口} / \text{投入} \\ &= \text{標準化効果} / (\text{投入} / \text{対象人口}) \end{aligned}$$

効率は、上記式の最後に示すとおり、対象人口一人当たりの投入に対する、標準化効果としても捉えることができる。

この手法を、フィリピンで展開するRHプログラムに適用して、評価を行うこととする。

フィリピンのリージョンでは、1992年から2002年までの11年間、JICAによるRH分野のプログラム協力が行われ、プロジェクト方式技術協力、青年海外協力隊、現地国内研修、開発福祉支援事業、無償資金協力事業等の異なるスキームが、リージョンにおけるRHの向上という共通の目標の下に展開された。本協力のプログラムアプローチに基づくロジックモデルは表3の通りである。

本プログラムの核となったプロジェクト方式技術協力では、最終的にリージョンの6州（人口779万人）に19のパイロットエリア（対象人口81万人）を設け、ここを拠点に活動を展開していた。

本プログラムの評価報告（駒澤2002、国際協力機構2004）によれば、2001年にパイロットエリア（直接裨益地域）とコントロールエリア（非裨益地域）において実施したサンプル調査において、次のようなプログラム目標に関連する指標の差異が見られた⁹。

	Pilot Area	Control Area
現在の子供の数	2.5人	3.2人
避妊実行率	69.8%	66.7%
子供の数の決定権が妻又は夫婦両方	95%	74%

上記から、現在の子供の数⁹、避妊実行率、子供の数の決定権が妻又は夫婦両方にある割合の3つの指標について、コントロールエリアと比較したパイロットエリアの指標の変化率を計算する。更にこれらが目標の達成を示す同等に重要な指標と仮定して、3グループの指標変化率の平均から

プログラムの効果を算出した結果は下記の通り。

$$\begin{aligned} \text{「子供の数」の指標変化率} &= 0.28 \\ \text{「避妊実行率」の指標変化率} &= 0.046 \\ \text{「妻又は夫婦による子供の数の決定割合」の指標変化率} &= \\ &0.283\% \\ \text{上記3つの指標変化率の平均から求めたプログラムの標準} \\ \text{化効果} &= 0.203\text{effect} \end{aligned}$$

このプログラムはUSAIDとの連携案件であったが、日本側が協力を受けた内容は主にUSAIDが開発した教材の活用であり、更にプログラム後半には連携はほとんど実施されなかった（国際協力機構2004、p.64）。このため、投入金額は、JICAの各種スキームの投入金額の合計から33.2億円と算定する。主な対象人口は、パイロットエリアの81万人であることから、インパクトと効率は下記のように算出される。

$$\begin{aligned} \text{プログラム協力のインパクト} &= 0.203\text{effect} \times 81\text{万人} = \\ &164,430 \text{ effect} \cdot \text{人} \\ \text{プログラム協力の効率} &= 0.203 \text{ effect} \times 81\text{万人} / 33,200\text{百万} \\ &\text{円} = 0.203 \text{ effect} / (33,200\text{百万円} / 81\text{万人}) \\ &= 0.203 \text{ effect} / (4.1\text{万円} / \text{人}) = 4.95 \text{ effect} \cdot \text{人} / \text{百万円} \end{aligned}$$

上記の効率を算出する中間式から、効率は、標準化効果（0.203 effect）を対象人口一人当たりの投入（4.1万円/人）で除した数値としても算出できることがわかる。

同じRH分野で活動する他のドナーやフィリピン政府のプログラム/プロジェクトについて、同じ指標群を使って測定すること、また、投入金額を共通の貨幣単位に換算することにより、効果、インパクト、効率の比較が可能となると考えられる。更に、フィリピン以外で展開しているRH分野のプログラム/プロジェクトについても、同じ指標、同じ貨幣単位で評価することにより、各国での協力の効果、インパクト、効率の横断的な比較が可能となると考えられる。

効果の差は、介入する側の条件だけではなく、介入される側の条件の違いによっても生じる。評価結果を比較することにより、効果的、効率的な介入のあり方についてのみならず、対象地域ごと、国ごとの促進要因、阻害要因についても、推定の糸口を与えられる。

最後に、この地域展開型案件のインパクト評価

表3 プログラムアプローチ・ロジック・モデル2(PLM2)概要表

対象地域：リージョンⅢ、対象裨益者：再生産年齢・既婚女性及び子供、協力期間：1992-2002年

上位目標	指標	プログラム目標	指標	成果	指標	投入された主な案件
リージョンⅢにおける、基礎的健康状態が向上する。	・乳児死亡率 ・5歳未満児死亡率	リージョンⅢにおけるRHが向上する。	・避妊実行率、避妊方法 ・現在の子供数、理想子供数 ・子供の数の決定者 ・母乳授乳実行率 ・産前・産後の検診回数	成果1：RH分野におけるサービスマスターの準備が整えられる。	・RHUの適切な配置 ・診察の待ち時間の短縮 ・RHUの設備の整備 ・RHUの職員数の確保 ・RHUの十分な投薬の確保	・プロ技フェーズI ・プロ技フェーズII ・開発福祉支援「地域保健強化計画(フィリピン小児病院)」 ・開発福祉支援「地域保健強化計画(Regina Carmeli 大学)」 ・無償資金協力 ・草の根無償資金協力 (8 案件)
				成果2：RH分野におけるサービスマスターの質が向上する。	・RHUの診察の適切さ ・RHUの治療の適切 ・RHUの職員による十分な情報提供	・プロ技フェーズI ・プロ技フェーズII ・現地国内研修 ・開発福祉支援「地域保健強化計画(フィリピン小児病院)」 ・開発福祉支援「地域保健強化計画(Regina Carmeli 大学)」 ・開発福祉支援「包括的リプロダクティブ・ヘルズ促進支援事業(PNGOC)」 ・草の根無償資金協力 (1 案件)
				成果3：RH分野におけるサービスマスターが提供される。	・出産立会人 ・母子手帳・成長記録表 ・使用率 ・IECへのアクセス	・プロ技フェーズI ・プロ技フェーズII ・青年海外協力隊 ・開発福祉支援「地域保健強化計画(フィリピン小児病院)」 ・開発福祉支援「地域保健強化計画(Regina Carmeli 大学)」 ・開発福祉支援「包括的リプロダクティブ・ヘルズ促進支援事業(PNGOC)」 ・草の根無償資金協力 (10 案件)

(出所) 国際協力機構『特定テーマ評価 - フィリピン人口・健康セクター / USAID連携 - Part 母子保健・家族計画(リプロダクティブ・ヘルズ)分野』p.55

と、第3節で述べた留学生政策のインパクト評価の異なる点を考察する。

留学生政策のインパクト評価においては、介入人口を留学生及び留学生による間接効果の及ぶ範囲の人々と捉えているのに対し、地域展開型案件のインパクト評価においては、介入人口を、直接、間接に効果が及ぶ者のみならず、効果が及ばない者も含めた、対象地域全体の人口と捉える点が異なっている。

留学生政策のインパクト測定においても、対象人口を対象国の国民全体と捉えることは、可能である。しかし、その場合には、留学生政策による直接、間接のインパクトの合計を国民人口で除した値が、国民一人当たりの平均変化率(政策効果)となるが¹⁰、国民人口の大きさから、実際に国民一人当たりの平均変化率を確認することは困難である。このため、元留学生と間接効果の及ぶ範囲の指標の変化率に着目して、インパクトや効率を評価することが現実的なアプローチである。

他方、地域展開型案件のインパクト測定においては、直接効果の及ぶ者、間接効果の及ぶ者、効果の及んでいない者に分けて指標を計測するよりは、サンプル調査等により、対象人口全体の指標を観測して効果を算出する方が、評価調査の手間も費用も少なく済む。

ただし、地域展開型案件の協力内容によって、介入人口の種類が限定されることはしばしば起こる。例えば家族計画・母子保健分野の協力においては、主な裨益対象者は再生産年齢にある女性と子供であることが多い。しかし、直接に裨益するのは女性や子供であっても、人口増加率の低減や母子保健指標の改善は、対象地域人口全体にとってプラスのインパクトを持つと考えられる。例えば、乳幼児死亡率の低下は、乳幼児の父親にとっても望ましいことである。

このことから、地域展開型案件の評価においては、直接裨益者を取り巻く人口コーホート(対象人口を直接裨益者数で除して、コーホートの平均人数を算出)が存在し、直接裨益者の指標の改善によって、コーホート全員、ひいては対象人口全員が裨益すると見なす。

この前提に立つことにより、地域展開型協力の評価において、対象人口の指標(群)の平均変化

率から標準化効果を算出し、対象人口を乗じてインパクトを求める本手法は、妥当性を有すると考えられる。

5. おわりに

本稿では、政策や援助などの介入について、目標の達成を端的に示す指標の変化に着目し、非介入人口と比較した介入人口の指標(群)の変化率、介入人口の規模、投入金額から、効果、インパクト、効率を測定する手法を提示した。更に、留学生政策と地域展開型プログラムの評価に、この手法を適用した事例を紹介した。

以上の評価結果から、本評価手法の利点として、次の点が挙げられる。

- (1) 非介入人口と比較した、介入人口における指標(群)の平均変化率を、介入の標準化効果としてeffectという単位で表現することにより、通常単位換算が難しい介入の効果を、定量的に表現することが可能となる。
- (2) 標準化効果に、介入人口を乗じることにより、介入がもたらしたプラスの変化の総量であるインパクトを、定量的に示すことが可能となる。このインパクトは、effect・人という単位で示すことができる。
- (3) 介入人口における変化の総量であるインパクトを、すべての投入を金額換算した投入金額で除することにより、単位投入金額あたりの効率を算出することができる。この効率は、介入人口一人当たりの投入に対する標準化効果とも捉えられる。効率の単位は、effect・人/金額単位として表すことができる。
- (4) 介入の効果を端的に示す指標(群)の選択、適切な非介入人口の選択、非介入人口と比較した介入人口における指標(群)変化率の適切な測定、介入人口の規模と投入金額の正確な把握、という条件がそろえば、政策、プログラム、プロジェクトのいずれの形態においても、本手法を用いて、標準化効果、インパクト、効率を定量的に評価することが可能となる。
- (5) 同じ指標(群)を使って、本手法による評価

を行う場合には、効果、インパクト、効率について、異なる介入主体間の比較や、異なる対象に対する比較、また、経年変化を調べることが可能となる。評価結果の差異は、主に介入する側と介入される側の条件の違いによって生じるため、評価結果を比較することにより、効果的、効率的な介入のあり方について、また、対象人口における促進要因、阻害要因について、考察する糸口が与えられる。

- (6) 本手法は、指標変化に着目する結果重視の評価手法の流れを汲むが、PEMなどのロジカルフレームワークを併用することにより、当該指標変化をもたらした因果関係の考察が容易となる。また、介入や評価の枠組みを、介入主体が、介入人口や関係者と共有することが可能となる。
- (7) 開発援助の評価は、主に開発の視点から行われてきたが、援助国に対する印象や理解に関する指標を設定し、非介入人口と比較した介入人口における指標の変化を測定することにより、援助国との友好や理解の促進といった、援助側の国益についても、定量的に評価することが可能である。

また、本評価手法の課題としては、次の点が挙げられる。

- (1) 目標の達成を端的に示す指標（群）の選定については、介入主体のみならず、介入対象人口及び関係者の意見を聴取して慎重に行う必要がある。
- (2) 指標には十分現れない公平性の観点や、副次的効果についても、十分配慮することが必要である。
- (3) 介入人口と同質の非介入人口（対照群）を選定することは容易ではないため、非介入人口と介入人口の差異により、指標の変化率に誤差が生じる可能性を十分考慮する必要がある。
- (4) 介入対象人口について、介入前と介入後の指標（群）の変化から、効果、インパクト、効率を測定する方法についても、検討する必要がある。
- (5) 複数の介入が並行して実施されている場合、

異なる開発介入がもたらす相互作用（例えば所得向上による保健指標の改善等）についても、検討する必要がある。

今後、これらの課題に取り組むことにより、本評価手法の改善を図り、適用範囲を広げたい。

注記

- 1 1993年～2000年のインドネシアからの留学生の平均在学年数が3.131年、タイからの留学生は3.482年であるが、これに日本語教育を日本で受ける期間を加え、約4年とした。
- 2 2001年にインドネシアで郵送と訪問調査により実施したアンケート調査では、日本の元国費留学生293名、元私費留学生37名、元米国留学生63名、非留学生60名の回答を収集した。2002年にタイで、郵送により実施したアンケート調査では国費165名、私費147名、米国246名、非留学生72名の回答を得た。（いずれも1954-1997年に留学または高等教育を開始して既に修了・卒業した者。日本留学生と米国留学生は、同窓会名簿から無作為抽出した者）。なお、母集団については、1954年から1997年までに来日したインドネシアからの留学生は4,964名（国費留学生1,742名、インドネシア政府派遣生1,810名、私費留学生1,412名）、また、タイからの留学生は5,007名（国費2,396名、タイ政府派遣生142名、私費2,469名）と推計される。日本留学生の回答者の属性と同窓会名簿の分析結果を照合した結果、アンケート回答者の属性は、同窓会名簿記載者の属性をほぼ反映していることが、筆者の先行研究（佐藤2002a、2002b、2003）で確認されている。
- 3 留学しなかった場合と比べた地位と収入の向上については、下記の設問（実際はインドネシア語、タイ語に翻訳）に対し、-20%、-10%、0、+10%、+20%、+30%、+40%、+50%、+60%、+70%、+80%の11の選択肢から回答を求めた。選択肢は、予備調査を踏まえて設定した。標準偏差は、収入の向上に関するインドネシアとタイの私費留学生の回答で、最も大きく、31.5%であった。
How much is your status up or down compared to the case you had not studied in Japan?

How much is your income up or down compared to the case you had not studied in Japan?

- 4 国費留学生と私費留学生の主な職業構成は下記のとおり (佐藤2004, p.44)
- <インドネシア> 国費: 大学教員58%、公務員・準公務員22%、企業勤務者14%
- 私費: 大学教員37%、公務員・準公務員17%、企業勤務者34%
- <タイ> 国費: 大学教員37%、公務員・準公務員18%、企業勤務者52%
- 私費: 大学教員14%、公務員・準公務員15%、企業勤務者66%
- 5 文部科学省データ、帰国留学生名簿、日本大使館資料から推計。
- 6 各国の国費留学生の予算は、各年の国費 = (留学生政策による直接・間接のインパクト / 国民人口) × 国民人口による直接・間接のインパクト予算額に、国費留学生に占める各国の国費留学生の割合を乗じ、また、私費予算額は、各年の私費予算額に、私費留学生に占める各国の私費留学生の割合を乗じ、それぞれ、2000年の物価指数を基準としてデフレートした上で、該当年度分を累計して算出した。また、授業料等収入については、国公私立大学の授業料、検定・入学料(4年に1度支払うとみなす)単価と私費留学生数、また授業料減免額(留学生予算に含まれない分)から、各年の私費留学生による授業料等収入を算出し、デフレートした上で、該当年度分を累計して算出した。
- 7 設問は、下記のように、1~5の選択肢から、文章の適合度を問う形となっていた。(実際はインドネシア語、タイ語に翻訳)
- To what extent, does the following statement apply (regarding the related topics)? Please circle the number.
- I like Japanese people.
- | | | | | |
|------------|---|---|---|-----------|
| 1 | 2 | 3 | 4 | 5 |
| Not at all | | | | Very much |
- 8 報告書に記載された目標に関連する指標から、目標の達成を最もよく示すと判断される指標を筆者が選択した。避妊などの家族計画に関する行動変化には通常時間がかかるため、これら指標はストック指標として扱った。
- 9 目標に照らし、少ない子供の数が望ましいため、指標の変化率算出にあたっては、逆数計算を行った。

10 国民全体に対するインパクトは、次のように算出され、留学生及び周辺の人に対する直接、間接のインパクトの合計と同じ値である。

$$\begin{aligned} & \text{国民一人当たりの指標の平均変化率} \times \text{国民人口} \\ & = (\text{留学生政策による直接、間接の指標変化} \\ & \quad \times \text{留学生数/国民人口}) \times \text{国民人口} \\ & = (\text{留学生政策による直接・間接のインパクト/国民人口}) \times \text{国民人口} \\ & = \text{留学生政策による直接、間接のインパクト} \end{aligned}$$

参考文献

- 小野達也 (2004) 「業績測定型の行政評価・政策評価における計量について - 目的達成度の把握を巡って - 」、『日本評価学会第5回全国大会要旨集録』、日本評価学会第5回全国大会実行委員会、47-50
- 菊田怜子・牟田博光 (2004) 「地域社会総合開発研修プログラムの総合的評価分析」、『日本評価研究』4(2):13-16
- 国際協力事業団 (1996) 『人作りプロジェクトのインパクト調査(フェーズ) - タイウボン職業訓練センター(UBISD)における事例研究』、国際協力事業団
- 国際開発高等教育機構 (1996) 『PCM手法に基づくモニタリング・評価』、国際開発高等教育機構
- 国際協力機構 『特定テーマ評価 - フィリピン人口・健康セクター / USAID連携 - Part 母子保健・家族計画(リプロダクティブ・ヘルス)分野』
- 国連開発計画 (1996) 『UNDP人間開発報告書 - 経済成長と人間開発』、国際協力出版会
- 駒澤牧子・駒澤大 (2002) 「協力プログラム評価の試み - JICAプロジェクト評価手法を適用して - 」、『日本評価研究』、2(2):101-114
- 佐藤由利子 (2002a) 「日本の留学生政策のインドネシアにおける影響 - 親日家養成の観点から - 」、『日本評価研究』、2(2):59-78
- 佐藤由利子 (2002b) 「日本の留学生政策のインドネシアにおける影響 - 人材養成の観点から - 」、『国際開発研究』、11(2):201-219
- 佐藤由利子 (2003) 「日本の留学生政策評価の試み - タイを事例として - 」、『留学生教育』、8:1-21
- 佐藤由利子 (2004) 「政策評価マトリックス (PEM) を使った定量的政策評価の事例 - インドネシアとタ

- イに対する日本の留学生政策評価』、『日本評価研究』
4(2):39-56
- 佐藤由利子(2005)『『開発』と『国益』、2つの視点からのODA評価の提案』、『援助と国益』、日本国際問題研究所、65-77
- 世界銀行(2003)『世界開発報告』、シュプリンガー・フェアラク東京株式会社
- 総務省(2005a)『地方財政白書』、総務省
- 総務省(2005b)『留学生受入れ推進施策に関する政策評価書』、総務省
- 田辺智子(2003)『プログラム評価の手法概観』、『政策研究』、16(5):22-27
- 統計情報研究開発センター(2004)『社会・人口統計体系都道府県基礎データファイル』、統計情報研究開発センター
- 中央教育審議会(2003)『新たな留学生政策の展開について - 留学生交流の拡大と質の向上を目指して - 』
- 三好皓一(2002)『プログラム・セオリー・マトリックスの活用について - 評価における分析の有用性を高めるために - 』、『日本評価研究』、2(1):11-27
- 三好皓一他(2003)『わが国評価における適切なプログラム・セオリーの構築を目指して - 国際協力評価と政策評価に焦点を当てて - 』、『日本評価研究』、3(1):43-56
- 牟田博光・菊田怜子(2005)『短期国際研修のインパクトに影響を与える要因の分析』、『国際開発学会第6回春季大会報告論文集』、国際開発学会第6回春季大会実行委員会、37-40
- 文部省(1970)『国費外国人留学制度の概要』
- 文部科学省(2004)『留学生受入の概況(平成16年版)』
- 山谷清志(1997)『政策評価の理論とその展開』、晃洋書房
- 龍慶昭、佐々木亮(2000)『『政策評価』の理論と技法』、多賀出版
- 留学交流事務研究会(2003)『留学生交流執務ハンドブック』、第一法規
- CIDA: Canadian International Development Agency (2000). *RBM Handbook on Developing Result Chains*. Gatineau: CIDA.
- Chen, H. T.(1990). *Theory-Driven Evaluations*. California: SAGE Publications.
- Knapp, P.V.D(2004). Theory-based Evaluation and Learning: Possibilities and Challenges. *Evaluation*, 10(1), 16-34.
- Kusek, J.Z. and Rist, R. C.(2004). *A Handbook for Development Practitioners: Ten Steps to a Results-Based Monitoring and Evaluation System*. Washington: World Bank.
- OECD(1991). *Principles for Evaluation of Development Assistance*. Paris: OECD.
- Rajan, R.G. and Subramanian, A.(2005). *IMF Working Paper: What Undermines Aid's Impact on Growth?:* Washington DC: International Monetary Fund.
- Sato, Y.(2005). A Case of New Policy Evaluation Utilizing a Logical Framework: Evaluation of Japan's Foreign Student Policy towards Thailand. *Evaluation*, 11(3), 351-378.
- Stame, N(2004). Theory-based Evaluation and Types of Complexity, *Evaluation* 10(1), 58-76.
- Schultz, T.W.(1963). *The Economic Value of Education*. New York: Columbia University Press.
- UNESCO(1963-1998). *Statistical Yearbook*. Paris: UNESCO Publishing & Bernan Press.
- Vedung, E(2000). *Public Policy and Program Evaluation*. New Jersey: Transaction Publishers.
- 国際協力機構「研修員受入れ事業とは」
<http://www.jica.go.jp/activities/ukeire/about.html>
visited in Jan. 2006.
- 国際協力機構「特定テーマ評価」
<http://www.jica.go.jp/evaluation/after/theme.html>
visited in Jan. 2006

(2006.1.30受理)

A Quantitative Evaluation Utilizing Standardized Effect Unit: Application to the Evaluation of Foreign Student Policy and Regional Cooperation Program

Yuriko Sato

Tokyo Institute of Technology
yusato@ryu.titech.ac.jp

Abstract

In this paper, the author proposes an evaluation method to measure the standardized effect by measuring the change of key indicators of the intervened group (in comparison with a non-intervened group). The standardized effect can be dealt as unit. She gives it a unit name ' effect '. Impact can be calculated by multiplying the standardized effect and the target population. Efficiency can be calculated by dividing this impact by the total input expressed in monetary unit. She shows two cases to which she applied this method: the evaluation of Japan's Foreign Student Policy and that of a regional Reproductive Health Program in the Philippines.

One of the merits of this method is to quantify the effect of intervention and to express it by a standardized unit ' effect '. By this quantification, it became possible to measure impact and efficiency in a standardized way.

By utilizing the same key indicators and conducting evaluation in this method, it will be possible to compare the standardized effect, impact and efficiency of different interventions. The difference of effect and efficiency will be caused by the conditions of the intervening side and the intervened side. So, this comparison will give clues to explore not only the effective and efficient way of intervention, but also the promoting and hindering factors in the target population.

Keywords

standardized effect, quantitative evaluation, efficiency, impact, result based evaluation

【研究論文】

情報流通に関わる政策の評価と分析 - エコマーク事業を事例にして

平原 隆史

京都大学大学院地球環境学舎

takashi.hirahara@k03.mbox.media.kyoto-u.ac.jp

要 約

少数政治主体による政策の立案や執行が難しくなり、取り組むべき政策課題が複雑になるにつれて、多種多様な政治主体が政策過程に参加し、政策関連の情報の収集や公開が行われるようになった。しかし情報を用いた政策や住民参加の政策の評価は、既存手法では十分な分析が出来ない。そこで政治学で定性的記述を行う政策ネットワーク論と、数理社会学で定量分析を行う社会ネットワーク分析を統合した政策ネットワーク分析を基に、情報政策や住民参加政策の評価法を考案した。本論文では環境政策の中でも情報的手法を用いた環境ラベルを例にとり、政策ネットワーク分析を行った。環境ラベルの基準形成過程を分析した結果、商品により政策主体の情報流通に果たす役割や、情報流通のネットワークの構造に違いがあることが明らかになり、それを基に政策提言や評価が行えるようになった。

キーワード

情報的手法、社会ネットワーク分析、政策ネットワーク、環境ラベル、政策評価

1. 研究の背景

政策決定への住民参加や政策に関わる情報公開・提供制度は、意思決定過程の透明性の確保、合意形成の円滑化、複雑な政策課題に関連する情報・知識・経験の集積と普及に役立つと考えられる。これら制度を導入する背景には、政策過程に参加できる利害関係者のみが政策関連情報を有し、参加できない人々が情報を持ってないという情報の非対称性によって生じるモラル・ハザードの防止、行政情報公開への市民運動による世論の圧力、政治課題の複雑化による旧来型行政システムによる行政サービス提供能力の低下の防止、など

が考えられる。

しかし、これら情報流通や住民参加の制度を分析できる手法は、限定的で少ない。現在使われている評価手法の多くは、制度導入の有無である。具体的な調査を例に挙げれば、国際的な調査としてOECDが2000年に加盟国の情報公開や住民参加について調査した報告書“Citizens as Partnership”(2001)や、日本国内の自治体では日経新聞と日経産業消費研究所が行っている「行政サービス・革新度調査」(1999, 2001)などに見られるように、多くの場合は当該制度や関連法の有無、もう少し詳細なもので制度の利用度、制度の認知度が評価の中心である。

OECD報告書の結果によれば、情報政策や住民参加制度は先進国といえども、情報公開や住民参加に関わるすべての制度が整っている国は少ない。そのため制度導入の有無が評価の中心課題となるのはやむを得ない。これは国内で同様の事例を扱った日経産業消費研究所の調査でも同様である。しかし、政策として情報を用いた手法が様々な政策領域で利用されはじめた以上、制度運用に関わる評価は、今後の制度の改善に不可欠である。

一方、政策手法には様々な形が存在するので、別の種類の政策評価手法が情報流通に関わる政策で利用可能かを考える。環境政策を例に政策手法を分類すれば、税、補助金、制裁金などを用いる経済的手法、危険物の使用制限、大気や水質汚濁など汚染原因物質の排出規制などを行う規制的手法が、代表的な政策手法として存在する。これは、食品安全や医薬品などの政策でも共通して用いられる一般的な政策手法である。しかし、地球温暖化対策のように、政策内容が多岐かつ総合的になると、経済的手法や規制的手法を補完するような政策手法が用いられるようになった。例えば、政策課題を利害関係者に周知させたり、問題解決に必要な情報を提供したりする情報的手法や、行政サービス需給の不均衡を是正するために、既存の行政組織以外の主体の参加を促す参加型手法と呼ばれるものがある。さらに政策評価を論じると、経済的手法には費用便益分析などの手法が存在し、規制的手法には規制影響分析といった手法が存在する。しかし情報型手法や参加型手法、とりわけ情報型手法において、これら分析手法では十分な政策評価が出来ない。情報普及や制度利用に掛かる費用を最小化することも重要であるが、政策導入の結果、情報が普及し、制度利用が促されたことを測定することが重要である。そこで制度の認知度や情報の普及度が、アンケート調査や特定期間での制度利用者数を基に測定される。先の日経産業消費研究所の調査にも、制度の利用度や情報の普及度の調査がある。しかし、それだけでは政策評価として不十分である。

一例として、日本のタイプ 型環境ラベルの「エコマーク」の認知度調査（日本環境協会エコマーク事務局、2001）を挙げる。エコマークの第三者認証機関である日本環境協会がエコマークに

関して2000年に行った調査では、認知度は全体の72%、エコマークをリサイクル商品と誤答した率は約2割（19.2%）、健康関連商品と誤答した率は1.4%、適正で安価な商品と誤答した率1.8%、また、第三者で認証していることや製品に基準が存在していることを知らなかったと回答した比率は7割以上であった。つまり、制度の存在は認知しているが、制度の内容までは理解していないということになる。特に、環境ラベルの場合、認定商品の利用が非認定商品に比較して、環境負荷を軽減するといったことを理解していないのでは、制度導入の目的を達成していないことになる。また2002年の調査でも、制度理解は進んだものの誤解もまだ多く存在している。

以上のように、認知度や情報の普及度だけでは不十分なケースがある。そのため、政策に関わる情報がどのように流通しているか考察することで、どこに正確な情報の流通の障害が存在するのかが分かる。これにより、情報発信と情報流通経路での問題点を洗い出すことが出来る。

また多くの場合、マスメディアは情報提供制度への関心は高くない。環境政策での情報流通制度である環境ラベルを事例に、マスメディアで情報提供制度がどの程度取り上げているのかを示す。日経gooのホームページ（<http://nikkei.goo.ne.jp/>）内に、日経4紙の記事索引データベースがある。そのデータベースで「環境ラベル」「エコマーク」「エコリーフ」²という見出し語で平成17年9月25日から過去1年間の記事数を検索した。検索結果は順に、18件、32件、7件であった。比較のため、同じデータベースで「環境問題」「地球温暖化」「気候変動」「公害」の見出し語を検索したところ、それぞれ順に680件、1383件、151件、537件であった。また、先述のエコマーク認知度調査でエコマークを知ったきっかけを調査した結果、実際に商品を見て知った回答者は約5割、マスメディアによって知った回答者は3割と、この調査でもマスメディアの果たす役割が小さいことが分かる。このように情報流通に関わる政策では、マスメディアの広報機能はあまり期待できない。結果、政策に関わる情報流通を把握するため、政策に参加するそれぞれの主体がどのような情報流通のネットワークを形成していくかを直接調査することが

必要となる。

では、政策過程で政策主体がどのようにネットワークを構築するのかを記述し、各主体がそのネットワークの中でどのような地位や役割を持つのかを分析するためには、どうしたらよいのか。

2. 政策ネットワーク分析

政策過程での各政策主体の参加に関わる構造を記述する方法として、定性的な手法としての「政策ネットワーク」と、定量的な手法としての「社会ネットワーク分析」がある。前者は政治学や行政学、後者は数理社会学を中心として、互いに独自の発展を遂げてきた。そこで筆者は、互いの特徴と共通点を基に1つの統合した分析手法として「政策ネットワーク分析」を提示しようと思う。

(1) 政策ネットワーク論

政治学や行政学の政策過程論において、意思決定に当たる主体は様々な形で分析がなされてきた。特に主体間の関係はモデル化により分析がなされてきたが、有名なモデルとして、アメリカにおける「鉄の三角形」モデルや、イギリスでの「ウェストミンスターモデル」などがある³。特にアメリカだけでなく日本でも用いられる「鉄の三角形」モデルは、政策決定に参加できる主体の数は少数に限られることが特徴である。

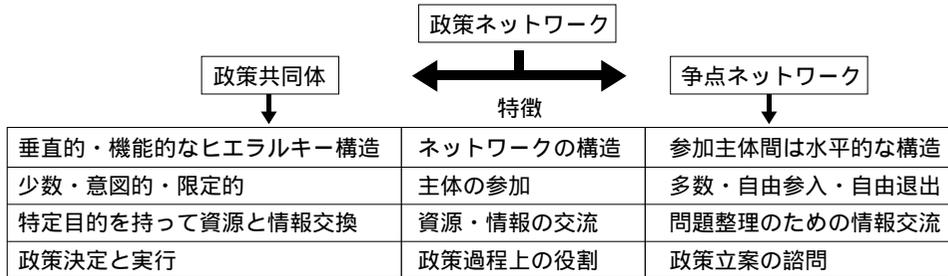
この少数主体のモデルは、政策課題が特定領域の特定問題を取り扱っているときには有効に機能したが、ある1つの政策課題や政策争点が複雑かつ複数の領域にまたがるような問題になると、少数政治主体では対処できなくなり多元主義的なアプローチが台頭してくる。アメリカの政治学者 Hecllo (1978) の議論はこの多元主義アプローチの代表例である。彼の議論を借りれば、アメリカでは1970年ごろまでに、政府の果たす役割は質的にも量的にも拡大した。例えば、年金、社会福祉、都市・住宅計画、公害をはじめとした環境問題など、その問題が複雑で、しかも外部経済性があるために利害関係者が無数にあるような問題では顕著となった。

こうした状況は、政策過程において新たな主体の参加を必要とし、問題解決のための専門家、住民運動をはじめとしたNPO、情報流通のためのメディア関係者などの主体が、新たに参加するようになった。こうした多種多様な主体による政策課題の問題解決や情報収集のために形成される各主体の関係と構造を、Heclloは「争点ネットワーク (Issue Network)」と呼んだ。

さらに、福祉国家による「大きな政府」と石油危機以降の経済不況は、先進国での財政危機を招いた。しかし1980年前後から、多くの先進国は、アメリカのレーガン政権、イギリスのサッチャー政権に代表される新保守主義的な「小さな政府」を目指した政権が成立することで、財政再建が行われた。これら政権での行政への姿勢は、行政の経済的効率化を目指すニュー・パブリック・マネージメントを中核とし、財政的にはケインズ主義からマネタリズム的手法、中央集権から地方分権への転換、公的部門の民営化への移行などの共通点が多い。またこれら政権は長期であったため、行政サービスは中央政府や省庁だけでなく、地方自治体、公企業、民間企業やNPOなどが多元的に供給する体制が確立された。加えて行政サービスの供給は、制度のもつ慣性や住民の要望もあり、削減が困難であった。結果、政策決定過程での多元的な参加主体体制が維持された。

一方、農業分野など主体の多元化が起らず、政策課題の構造や政策参加主体がほとんど変化しない政策領域もある。これら領域では「鉄の三角形」モデルのような少数政策参加主体の構造が維持されやすい。この少数主体が特定政策領域を支配するモデルを「政策共同体 (Policy Community)」と呼び、「争点ネットワーク」の対極に置かれた。またこれら2極の連続体の間に、現実の主体の政策参加構造は位置しているとする、イギリスの Rhodes と Marsh (1992) らによる「政策ネットワーク (Policy Network)」論が発表された。図1にまとめたように、政策ネットワーク論は各政策主体の政策過程参加構造の特徴を分類し、定性的な記述を可能にした。

図1 政策ネットワーク概念図



(出所) 正木(1999)の整理をもとに筆者作成

(2) 社会ネットワーク分析

この政治学や行政学の研究の発展とは別に、数理社会学のソシオメトリーを基に、数学のグラフ理論やコンピュータ科学の進展を吸収しながら、個人間、組織間、個人・組織間などの社会的関係を数理的に構造化して分析しようとする「社会ネットワーク分析」という立場がある。

簡単にこの分析の特徴を言えば、個人や法人・組織などを主体として点、その主体間の関係を線や矢印で捉え、表記することにある。これら点や線・矢印からなるグラフをネットワークとし、ネットワークを社会集団や社会構造と捉える。逆に言えば、グラフの中の位置で、点の役割を数理的に割り出せる。さらに社会構造の中で、主体がどのような役割を果たしているか数理的に分析できる。この手法をもとに、1970年代以降様々な研究がなされた。その中には、政策主体の関係性を分析する研究もある。

特に情報政策や住民参加を考えれば、政策主体間のコミュニケーションの構造を分析することにより、政策に関わる情報の交換構造を把握出来る。加えて、政策に関わる合意形成でどのような障害が存在するのかなど、政策過程に参加する主体間の関係が、分析できるようになる。

この社会ネットワーク分析の中で、主体の地位や役割を分析する手法は、中心性分析である。中心性分析によりネットワーク中での優位的な立場を把握し指標化できる。

本論文では、多数ある中心性の中から、特に次数中心性と近接中心性を用いて主体間の関係を分析する。

次数中心性

次数中心性も近接中心性も、先行する社会ネットワーク研究を基にFreeman(1979)が分析手法としてまとめたものである。そこで、2つの中心性について説明する。

先に、次数中心性(Degree Centrality)から説明する。公式は次の通りである。

$$C_d(v_i) = \frac{\text{deg}(v_i)}{n-1} \dots$$

点 v_i がネットワーク上の点 v_j を除く任意の点と接している数が次数 $\text{deg}(v_i)$ である。ネットワーク上の点の数を n とすると、次数の最大は $n-1$ となる。よって、最大の次数を分母とし、任意の点 v_i の次数 $\text{deg}(v_i)$ を分子とすれば、次数中心値 $C_d(v_i)$ が求められる。中心値は0から1の間の値をとり、0に近いほど他者との関係が希薄、1に近いほど他者との関係が稠密であることを表す。ゆえに中心値の結果は、局所ネットワークでの主体の役割の重要度を示す。つまり、派閥などでの主体の役割を導き出す時に用いられる。

近接中心性

同様に、近接中心性(Closeness Centrality)の公式は次のようになる。

$$C_c(v_i) = \left(\frac{\sum (v_j)}{n-1} \right)^{-1} = \frac{n-1}{\sum (v_j)} \dots$$

グラフ上の点 v_i から他の点までの距離の総和をステータス $\sum (v_j)$ と呼ぶ。この値が小さくなればなるほど、情報・資源の伝達の経路が小さくなり、ネットワークの中心に位置することがわかる。

よって、近接中心値 $C_c(v_i)$ はステータスの逆数として定義し、ステータスが最小となるのは $n-1$ であるので、これを分子として、ステータスを分母にとれば、次数中心値と同様、値は0から1の間をとり、1に近いほど全体への情報・資源の伝達能力が高いことを示す。つまり、ネットワーク全体での主体の役割の重要度を示している。

本論文のように情報流通に関わる政策を分析するならば、情報交換関係、つまり各主体間のコミュニケーション関係の有無について中心性分析を行えばよい。

(3) 政策ネットワーク分析

政策ネットワークと社会ネットワークを概観したが、この2つの研究の流れは、それぞれの研究を軽視していたわけではない。特に、政策ネットワークの研究では、数理社会学の研究に注意を払ってきた。

政策ネットワーク論には、Dowding (1995) 論文が契機となった「メタファー」論争がある。DowdingはRhodesとMarshらによる政策ネットワーク論が記述のメタファーに過ぎないという批判を行った。それに対しMarsh (1998) はその批判に対し、数理化による比較可能な記述法の確立に重要性があることを認め、数理社会学のネットワーク概念も条件が整えば、受け入れるべきであるとDowdingの批判を受け入れた。またRhodes門下の研究者Hugh (2000) は、ネットワーク図を描きながらイギリス財政での、政策主体の役割の変化を時系列的に追っている。

一方、数理社会学を用いた政策ネットワーク研究では、Knockeや辻中 (1996) の日米独の労働政策での比較があるが、こちらでは政治学の政策ネットワーク論の存在に注意を払っていない。また、Marsh (1998) はこうした数理社会学の政策研究は、ネットワークの主体の役割にのみ焦点を当てて、ネットワークの構造評価が出来ていないと論じている。

では政策ネットワーク自体は、数理的に説明できないのか。筆者は、次のように考えることで説明可能であると考えます。

図1で示したように、「政策共同体」は少数主体、

ネットワークにヒエラルキーがあるという特徴があり、「争点ネットワーク」は多数主体で、ネットワークは水平的であるということである。これを中心性と絡めて考えれば各主体の中心性の値が分散するのが「政策共同体」、分散しないのが「争点ネットワーク」となる。つまり、中心性とりわけ近接中心性の分散で、ネットワークは水平構造か垂直構造かが分かる。さらに次数中心性を見ることにより、派閥でのヒエラルキーの存在も分かる。加えて、「争点ネットワーク」として機能するには各主体の中心性の平均が高い値でないとは有効に機能していないということもわかる。以上、筆者は2つのアプローチは統合しながら、政策ネットワーク自身の数理的説明も可能であると考えます。また、政治学と数理社会学の2つのアプローチを、中心性概念を基に統合した分析法を、ここではMarsh (1998) の言葉を借りて「政策ネットワーク分析」と呼ぶことにする。そこで政策ネットワーク分析から見た「政策共同体 (略記：政策)」と「争点ネットワーク (略記：争点)」は表1のようになる。

表1 政策ネットワーク分析の特徴

中心性の特徴	政策	争点
主体の数	少ない	多い
近接中心性の分散	大きい	小さい
次数中心性の分散	大きい	小さい
近接中心性の平均値	無関係	高い

(出所) 正木 (1999) の議論をもとに筆者作成

また情報的手法を用いた政策や住民参加などの参加型手法を用いた政策で考えれば、「争点ネットワーク」的な政策ネットワーク構造になっていることが、政策が機能するために重要である。そこで、環境ラベルを基に、政策過程のネットワークと、政策主体の役割について、議論の検証のために実証分析を行う。

3. 事例分析：エコマーク

実証分析の対象に環境ラベルを事例に選んだ理由は、環境問題における情報流通を目的とした政

策だからである。また、意思決定に関わる主体の多様性が制度上確保されることも重要である。これによりネットワークの性格を論じることが出来る。もし主体の属性が2~3くらいに分類できるとすれば、政策ネットワーク分析より、既存のゲーム理論や費用便益分析の方が容易に分析できる。また、主体の属性の数があまり多くないことも重要である。過剰な主体の属性数は、ネットワークがあまり詳細すぎて、全体の構成を把握しにくくなる。よって研究対象が、適当な主体の属性に分けることが可能な場合に、この分析は最も効力を発揮すると思われる。

こうした分析上の特徴に、環境ラベル、とりわけ日本のエコラベルの基準形成過程は当てはまっている。

(1) エコマークと基準形成の過程

エコマークは、日本の第三者認証型環境ラベルで、ドイツについて世界で2番目に1989年に導入された。国際標準化に際し、1996年以降、エコマーク商品の認定基準は、生産から廃棄までのすべての過程で出来る限り環境に配慮するような基準に変更された。現在45商品類型が存在するが、すでに34類型が新基準の類型となっている。また、この類型は最低でも5年に1度、類型そのものと基準の見直しが行われる。本論文ではこの基準形成作業での政策ネットワークを分析する。

最終的な政策決定と審議人員はそれぞれ、エコマーク運営委員会やエコマーク類型基準委員会が決定するが、実際の基準決定作業に携わるのは、作業部会（以下WG）である。WGは1つのタイプの基準形成作業に10名が参加し、生産者側（企業、生産者団体など）から3名、消費者側（NPOなど）から3名、中立側から4名（研究者、行政関係者）が参加するのが一般的である。WGによって若干の差があるが、基本的に生産側と消費側の人数は同数となるよう日本環境協会のガイドラインで規定されている。さらにWGで決定された基準は、WG外部の一般の人間に対して、意見を公聴するパブリックコメントの期間が60日ある。この手直しを経て、エコマーク運営委員会で基準は了承される。よってWGによる基準形成作業は、各種の

属性を有するWGと日本環境協会の担当者やエコマークに関わる各種委員、さらにWGの外部には認定商品に関わる企業や、一般消費者などの主体が存在し、多様な属性の主体が確保される。一方、主体の属性は多くても30は超えることはない。

また先述の通り、環境ラベルではマスメディアが機能しないので、こうした基準形成作業期間での各主体間の情報交換以外に、政策に関わる正確な情報が一般に広まる機会はほとんど存在しない。これはマスコミのアジェンダ設定機能に関わらない政策であれば、どんな政策でも同様なことがいえる。

(2) 事例分析：エコラベルに関わる政策ネットワークと情報流通構造

エコマークを分析するに当たり、エコマークの事務に当たる日本環境協会の協力を得てWG成員に関して、予備調査と本調査のアンケートを行った。予備調査は、2004年での紙関連商品でのWGでの意思決定に関わる情報流通ネットワークについてアンケートを行った。本調査は2005年の8月にこの3年間にWGに関わった成員全員に対して、アンケートを行った。

予備調査では、政策主体間のコミュニケーション構造が不明であるので、この構造がどのようになっているかを調べる必要があった。そこで政策策定過程の前後の期間を含めた、電子メール、電話、ファックス、郵便書簡、直接面談の回数や頻度を調べ、どういったコミュニケーション経路と構造が出来上がっているかなどの詳細なデータを、アンケートを通じて集めた。また将来的に、ある政策に関わるイベントが起こる事前、進行中、事後において、政策主体間のコミュニケーション・ネットワークの構造が変化するか、情報流通の中心性が変化するかなど、比較静的な構造変化の測定を念頭において調査を進めた。

また予備調査の対象となった商品はどれも紙関連商品であり、それぞれ「紙製の事務用品」「包装用の用紙」「紙製の包装用材」となっている。この3つにかかわるWGは先にも述べたが10名で、内訳は供給者代表・消費者代表が各3名、行政・専門家などの中立的な立場から4名である。また、

WGは3つとも2003年の2月から12月までで9回行われ、2004年3月1日に改定基準案が公表され、60日間公告され³、これを受け、基準の最終決定を行う類型・基準制定委員会で改正案が既に了承されている（2005年9月現在）。

調査は日本環境協会の協力を得て、全WG成員と協会側の事務担当者、情報流通に関してアンケートを行った。期間は2004年7月13日から31日まで、郵送法で回答数は4件（内訳：供給・中立各1、消費2）と、事務担当者に対しては、主体別に諮問や陳情を受けた回数について調査した。また調査で用いた主体の分類は、次の通りである。WG帰属の有無：2通り×（生産者（団体）・消費者（団体）・環境団体・行政関係者・研究者の6分類）、日本環境協会の事務担当者、WG以外のエコマーク委員の計14通りに分けて、時間軸（WG開始前・後）・コミュニケーション回路（電子メール・電話とファックス・書簡・直接面談）・情報内容別（基準に関わる情報収集・基準変更のための交渉）に分けて、情報のやり取りの回数に関して質問をした。

予備調査の結果を要約すれば、コミュニケーション回路や情報内容とは無関係で時間軸だけがネットワークの性格を変化させる。WG会合が始ま

る前の各アクター間のコミュニケーションはほとんど皆無であり、WGグループが始まるとコミュニケーションは活発化する。しかし、同じ属性もしくは日本環境協会の事務担当者の間は活発化するが、異なる属性での会合以外のコミュニケーションは基本的にほとんどなされない。とりわけ、消費者団体のWGは会合以外の他のアクターとの交渉は望まない旨の回答があった。

さらにアンケート結果は、Excelで集計して関係行列を作成し、UNINET6.0により中心性を計算したのが表2である。結果から分かるように、情報流通の中心は日本環境協会の担当者であり、WGに属する成員であることが分かる。結果として、情報はWGの外部にはあまり流れてないことが推測できる。また基準の内容について一般の消費者が知る情報は少なく、アンケートからも生産側とは直接のコンタクトはないので、入手した情報も消費者団体や環境団体からのバイアスがあっても補正しにくく、正確な理解が難しい状況にあると考えられる。また、各主体に属する数字は異なる。WG以外のエコマーク委員は30名を超えることはなく、各WGは4人以上になることはない。一方、消費者や企業や生産者は無数にいる。そのため、消費者の中心性が高くても参加している人

表2 WG期間でのコミュニケーション・ネットワークの中心性分析の結果（予備調査）

主体	回数	ステータス	回数中心値	近接中心値
WG以外のエコマーク委員	1	25	0.0769	0.3600
エコマーク事務局関係者	7	13	0.5385	0.6923
生産者側のWG関係者	2	17	0.1538	0.5294
WG以外の生産者団体	0	N.A	0.0000	0.0000
WG以外の生産者（企業等）	1	19	0.0769	0.4737
消費者側のWG関係者	2	17	0.1538	0.5294
WG以外の消費者団体	0	N.A	0.0000	0.0000
WG以外の消費者	2	20	0.1538	0.4500
環境団体のWG関係者	2	17	0.1538	0.5294
WG以外の環境団体	0	N.A	0.0000	0.0000
行政のWG関係者	1	19	0.0769	0.4737
WG以外の行政関係者	0	N.A	0.0000	0.0000
WGの研究者	2	18	0.1538	0.5000
WG以外の研究者	2	18	0.1538	0.5000

（出所）筆者作成

の数を考えれば、ほとんど流通していないのも同様である。加えて、この予備調査からいろいろ問題点が浮上してきた。研究者のWGがないため正確な構造とは言い切れない。また、環境団体と消費者団体の区別は実際のところ難しく、予備調査ではこの2つの主体の情報流通の総量を2つに分けて分析を行った。加えて、情報の照会と基準形成活動は分離しにくい面もあるので回答に苦慮するという意見もある。エコマークに関わるコミュニケーションの頻度を有無に変えたとしてもグラフの構造は変化しないので、簡略化した設問が必要である。また、企業でも流通にかかわる主体の働きなど、現実には様々な問題がある。そこで、こうした問題点を踏まえて、本調査を行った。

予備調査で挙げた欠点を補うために、2005年8月にこの3年間にWGに在籍をしていたすべての人に対して、より簡便な形で郵送回答法によるアンケートを行った。

この調査の対象になった総数は132名、2005年9月10日までの回答数は69名（うち有効回答数64）である。また、1人で複数のWGに参加するのでのべ表記で表すと、のべ220名中106名（うち有効のべ回答数99）から回答があった。今回対象になった商品類型は22商品にわたり、内容も複写機や時計、プリンターなど一般消費者も多く使うものや、生分解性潤滑油や土木製品のような法人向けの商品もあり、より一般的な構造について考察できるようになった。

アンケートで質問した項目は、自分の関わった商品類型と自分の立場（消費・生産・研究・行政・その他）を回答してもらった上で、14の主体に対してのWG開催中のエコマークに関わるコミュニケーションを5段階の頻度と、同じ分類の主体に対するWG以外の期間でのエコマークに関わるコミュニケーションの有無を質問した。ちなみに14の主体は、自分の参加するWGに属していないエコマーク委員、環境協会の担当者、WGでは生産者代表、消費側代表、行政代表、研究者代表、流通者代表（存在している場合のみ）の5つ、WGに属さない主体は、認定商品に関わる製造業者、認定商品に関わらない製造業者、流通業者、NPO（環境問題と消費問題に関わるもの）、一般消費者、研究者、行政関係者の7つの計14主体に

分けてアンケートをとった。また、予備調査で主体間のコミュニケーション頻度の正確なデータは、多くの場合残ってないことが分かった。そこで、回答者の記憶に頼る次の5段階の主観尺度に分け、正確なコミュニケーション頻度の代替とした。情報交換がない、WGでのみ情報のやり取りがある（WG以外の主体では、まれに情報のやり取りがある）、どちらかと言えば相手からの情報の流通が多い、どちらかと言えば自分の方から積極的に情報を発信する、相互に活発な情報のやり取りをしている。この5段階に分けたのは、今後の研究で、主体間での相互情報の非対称性にも考慮した分析を想定しているからである。

このアンケート結果から、WG生産者代表・WG消費者代表・WG研究者代表・WG行政代表のそれぞれののべ数を重み付けして1人当たりの平均スコアを出した⁵。次にインフォーマルな関係性を導き出すためWGに関わる主体と環境協会の担当者はスコアから2点を引き⁶、それ以外の主体からは1点を引いた。結果、得られた点数を四捨五入して1点以上ある場合、主体間はエコマークに関わるコミュニケーション関係を持つ、0点以下は関係を持たないということにした。また、WG以外の主体間のエコマークに関わるコミュニケーションが日常的に行われることは、経験に照らして考えにくい。そのため、WG以外に属するアクター間のエコマークでのコミュニケーション関係は存在しないと仮定した⁷。この平均値から導き出したものを、以下平均モデルと呼ぶことにする。また、各種商品類型の分析では、アンケート結果がないものは平均モデルを基に、アンケート結果のある部分だけWG1人当たりの関係性に転換して、関係行列を構築した。

またアンケートの結果から、流通関係の主体とはWGの有無を含めて存在していないので分析から外して、13の主体の属性間で次数中心性と近接中心性を測定した。その13の主体は次の通り、自分の参加するWGに属していないエコマーク委員（略記号a、以下同様）環境協会の担当者（b）、WGでは生産者代表（c）、消費側代表（d）、行政代表（e）、研究者代表（f）、WGではない主体からは、認定商品の製造業者（g）、認定外商品の製

造業者 (h)、流通業者 (i)、環境・消費関連NPO (j)、一般消費者 (k)、研究者 (l)、行政 (m) となる。

最後に当論文での分析対象商品は、生分解性潤滑油 (以下、潤滑油、類型番号110)、紙製の事務用品 (以下、紙製品、同112)、印刷インキ (以下、インク、同102)、時計 (類型番号134) とした。これらの商品はアンケートの回答率が高く (潤滑油7名、潤滑油以外6名) で、統計的に考えて結果の信憑性が高いことが選定理由である。また、それぞれの商品は特性がある。潤滑油とインキは消費者が法人利用中心で限定的であるが、紙製品と時計は法人利用だけでなく一般消費者の利用が中心である。とりわけ、紙製品に関しては先行する予備調査と比較可能である上、日本環境協会の調査 (2003,2005) でも一般消費者の利用経験が上位5位以内に位置する商品である (2002年末では3位、2004年末では4位)⁸。また時計に関しては、まだ導入されていない現在パブリックコメント募集中の類型であり、こうした特徴が、コミュニケーション・ネットワークにどのような影響を与えるかを考えた上である。これら商品の特性や基準形成の特性のネットワークに与える影響を、政策ネットワーク分析を行って論じる。

(3) 分析結果

まず分析結果から、平均モデルでの主体のネットワークについて図示する (図2)。

図2はExcelで関係行列を作った後、Netdraw2.8によってネットワーク図を作成したものである。また記号は、先述の主体の属性の略記号と同じである。

図2から、平均的なエコマークの意思決定はWG中心、とりわけ環境協会の担当者中心に、比較的稠密な結びつきで主体間のコミュニケーション・ネットワークが形成されていることがわかる。

また各種の中心性分析は、表2同様にUNINET 6.0で分析した結果、表3ようになる。主体のアルファベットは、同じく先述の略号と対応している。

さらに表1に従い、Excelで政策ネットワーク分

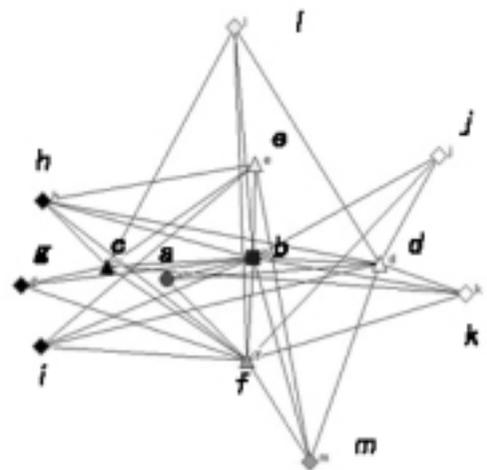
析を行った結果が、表4ようになる。

また政策ネットワークに関わる各種中心性の分散分析の結果は変動係数で出している。変動係数は標準偏差を平均で割るので、データの内容に関わらず相対的な比較を可能にする。この利点を踏まえ、政策ネットワーク分析では、変動係数を利用した。

分析により、予備調査の項目を除いて、近接中心性の変動係数は0.2前後の結果を得た。加えて、係数から分散の幅は小さいと判断でき、全体の情報流通に関しては争点ネットワーク的であると考えられる。また、近接中心性の平均値も0.6前後と高く、争点ネットワーク的な特徴を示している。一方、次数中心性は変動の幅が大きい。表3よりWGとWG内外で各主体の結果の差が大きい。つまりWG内外で中心 - 周縁関係があると考えられる。また予備調査に関しては、明らかに本調査の各種結果と差が大きく、主体の分類の仕方やアンケートの数など、検証すべき点がある。この問題については後述する。

個別に見ていくと、時計はまだ決定されていない類型である事が影響して、近接中心性が低く、また変動の幅が大きい。つまり政策共同体的な要素が強いことが分かる。新基準形成を目的にしているため、意見収集や諮問より、実行的に意見を統一しようとする傾向が各主体に共有されている

図2 平均モデルのネットワーク図



(出所) 筆者作成

表3 基準形成での主体の役割

(a) 次数中心性					
主体	次数中心値				
	平均	潤滑油	紙製品	インク	時計
a	0.25	0.17	0.25	0.17	0.08
b	1	1	1	0.91	1
c	0.58	0.58	0.42	0.25	0.25
d	0.67	0.75	0.75	0.75	0.08
e	0.42	0.5	0.25	0.41	0.42
f	0.83	0.58	0.83	0.08	0.17
g	0.25	0.33	0.25	0.25	0.17
h	0.42	0.33	0.25	0.33	0.25
i	0.33	0.17	0.25	0.25	0.17
j	0.25	0.17	0.25	0.17	0.08
k	0.33	0.25	0.25	0.17	0.08
l	0.33	0.33	0.25	0.17	0.08
m	0.33	0.33	0.33	0.25	0.17

(b) 近接中心性					
主体	近接中心値				
	平均	潤滑油	紙製品	インク	時計
a	0.57	0.55	0.57	0.55	0.52
b	1	1	1	0.92	1
c	0.71	0.71	0.63	0.57	0.57
d	0.75	0.8	0.8	0.75	0.52
e	0.63	0.67	0.57	0.63	0.63
f	0.86	0.71	0.86	0.5	0.55
g	0.57	0.6	0.57	0.57	0.55
h	0.63	0.6	0.57	0.6	0.57
i	0.6	0.55	0.57	0.57	0.55
j	0.57	0.55	0.57	0.55	0.52
k	0.6	0.57	0.57	0.55	0.52
l	0.6	0.6	0.57	0.55	0.52
m	0.6	0.6	0.6	0.57	0.55

(出所) 筆者作成

可能性が高い。また、潤滑油、紙製品、インクに関しては既存の基準を改定する作業であるため、より基準の実行性を高めるために広く意見収集をしようという傾向、つまり争点ネットワーク的な性格に傾いたのではないかと推測される。

表4 基準形成の政策ネットワーク分析

	平均	潤滑	紙	予備	インク	時計
主体数	13	13	13	14	13	13
次数変動	0.52	0.59	0.65	1.14	0.76	1.08
近接変動	0.20	0.20	0.21	0.68	0.18	0.22
近接平均	0.67	0.65	0.65	0.36	0.61	0.58

(注) 変動は変動係数、潤滑は潤滑油、紙は紙製品、予備は予備調査の結果の略

(出所) 筆者作成

4. 問題点と今後の研究

(1) 問題点

先にも述べたが、同じ類型で結果が違うのは、アンケートの結果が問題なのか主体の属性の分類の違いなのかを考える必要がある。

そこで、2つの調査で共通点のある主体を統合し、11主体(エコマーク委員、環境協会の担当者、WGでは生産者代表、消費側代表、行政代表、研究者代表、WGではない主体からは、生産者、環境・消費関連NPO、一般消費者、研究者、行政)で中心性の数値を相関分析した。

まず予備調査による紙系商品類型の関係モデル(以下、紙1モデル)平均モデルを基に紙関係類型に関わるアンケート結果の回答がある部分は回答に置き換えるモデル(以下、紙2モデル)紙2モデル同様、予備調査で完全回答がある部分だけは置き換えるモデル(以下、紙3モデル)とした。それぞれUninet6.0で中心性を分析し、Excelで中心性の結果の相関分析を行った。

結果は表5ようになる。近接中心性と次数中心性の双方で、平均モデルを基に作成した紙3モデルよりも、予備調査の結果のみで作った紙1モデルは相関性が高い。さらに次数中心性に着目すれば、すべての紙モデルのうちで最も平均モデルに近いモデルは紙1モデルである。また次数中心性の方が近接中心性より相関性が高いが、どちらの中心性も相関の値は低くない。よって標本数の差よりも、主体の属性の分類差が中心性に影響を与えることが明らかになった。

表5 モデルによる中心性の相関性比較

(a) 次数中心性				
	平均	紙 1	紙 2	紙 3
平均	1			
紙 1	0.8	1		
紙 2	0.78	0.91	1	
紙 3	0.78	0.98	0.91	1
(b) 近接中心性				
	平均	紙 1	紙 2	紙 3
平均	1			
紙 1	0.52	1		
紙 2	0.76	0.62	1	
紙 3	0.5	1	0.61	1

(出所) 筆者作成

(2) 今後の研究

以上のように中心性を用いた分析は、政策に関わるコミュニケーション・ネットワークの構造を明確にし、その問題点や障害を明確にする。

しかし、いくつかの課題が残されている。

本論文では、情報流通量の差異を投影した分析を行っていない。これはグラフ理論では有価性と呼ぶが、関係の有無だけでなく、主体の役割の差や取引量・交換量を投影する手法である。筆者は、この有価性に対して拡張近接中心性という議論(平原2006)から、予備調査での分析を試みている。この有価性のもつ関係の分析を、本調査でも行う必要がある。有価性を考慮しなければ、コミュニケーション頻度を踏まえた分析は不可能である。また、既存の関係の有価性に対する分析手法としてボナッチ中心性や情報中心性などの方法があり、これら既存の方法との分析の優位性と問題点を明らかにする必要がある。

また、他の商品類型での分析を行い、さらなる比較分析によって政策評価手法としての洗練をはかる必要がある。

最後に、WG前後でのアンケートを行っているため、比較静的な分析も行う必要がある。これら3つの問題を今後の課題とし、研究の深化と情報的手法の分析手法としての政策ネットワーク分析を洗練させることを目標として、本論文の結語

としたい。

謝辞

今回の研究において、京都大学大学院地球環境学堂の松下和夫教授、小畑史子助教授、松本泰子助教授をはじめ、研究室や研究科の教官・同僚に研究上様々なアドバイスを頂いた。また、アンケート準備において、日本環境協会の田口整司エコマーク事務局長をはじめ、その他多くのエコマーク事務局の方々に御協力頂いた。またアンケートに回答を頂いた各WGグループの方々に、この紙面を借りて感謝したい。

注記

- 1 環境ラベルは、その特徴によって国際標準(ISO14020)で3つに分類されている。ラベルを第三者が認証するタイプ、自己宣言するタイプ、数値明示するタイプがある。タイプは、認定商品のライフサイクルも含めて環境負荷を低減することが目標とされており、国際的に最も一般的に用いられている。
- 2 「鉄の三角形」の主体はアメリカの場合、議会委員会(小委員会)・関連行政機関・利益団体を指し、「ウェストミンスターモデル」では、国会・中央政府指導による議会での多数決による意思決定を指す。
- 3 日本環境協会によれば、この3つのパブリックコメントはそれぞれ「紙製の事務用品」が163件(うち非生産者が2件)、「包装用の用紙」が13件(うち非生産者が1件)、「紙製の包装用材」が6件(非生産者からはなし)寄せられた。
- 4 WG以前のコミュニケーションは、行政関連のWGからの情報照会が5件、企業からの陳情が11件、環境協会の別のWG担当者の陳情が3件、生産者WG同士の話し合いが年に数回(1ヶ月に一回以下)、行政WGがWG以外の行政関係者と話し合うのが年に数回、という回答があった。一方、WGに入ると頻繁なところで1日1回程度、頻度の低いところでも月1回前後のやり取りが発生する。
- 5 アンケートの結果からはその他に分類される結果は出なかった。その他に回答したものでも、立場を書

- かせるようにしたので、その回答によって4つの立場に分類することができた。
- 6 頻度で示したがWGの2点はWG以外でのインフォーマルなコミュニケーションはないことになる。それに対しWG以外の主体では2点以上はコミュニケーションがあることになる。
- 7 厳密さを問えば、わずかにはエコマークに関わるコミュニケーションは存在するであろうと想定される。しかし、先の日本環境協会の調査でも、多くの場合、一般消費者がエコマークを知るのは実際に商品を見てであり（全体の5割程度）、友人などの紹介（全体の1割程度）ではないことが示されている。このように、エコマークのコミュニケーションはほとんど存在していない。しかも意思決定に関わるコミュニケーションとなると、ほとんど皆無といってよい。
- 8 2002年末の調査だけが利用経験を聞いているが、その利用率は約8割である。また2004年調査では、現在家にある商品として調査していて、9%の家庭に存在している。ちなみに両調査共に順位の高い商品は再生プラスチック利用商品（2002年調査1位、2004年調査2位）と衛生用紙（トイレットペーパー、同2002年2位、2004年1位）である。

参考文献

- 金光淳（2003）『社会ネットワーク分析の基礎』、頸草書房
- 鈴木義一郎（1998）『現代統計学小事典（講談社ブルーバックスB1208）』、講談社
- 日本環境協会エコマーク事務局（2001）「エコマークガイドライン・規定集」、日本環境協会
- 日本環境協会エコマーク事務局（2001）「エコマークと消費者意識調査」、日本環境協会
- 日本環境協会エコマーク事務局（2003）「第2回エコマーク商品の消費者モニタリング調査および認知度・信頼性調査」、日本環境協会
- 日本環境協会エコマーク事務局（2005）「第3回エコマーク商品の消費者モニタリング調査および認知度・信頼性調査」、日本環境協会
- 日本経済新聞社、日経産業消費研究所編（1999）『住民サービスここが一番：全国都市番付』、日本経済新聞社
- 日本経済新聞社、日経産業消費研究所編（2001）『全国住民サービス番付』、日本経済新聞社
- 平原隆史（2006）「政策ネットワークと中心性分析の関係性の考察」、『情報処理学会誌』47巻3号、近刊
- 正木 卓（1999）「政策ネットワークの枠組み 構造・類型・マネジメント」、『同志社政策科学研究』創刊号、91-110.
- Heclro, H. (1978). Issue Networks and the Executive Establishment. In A. King (Eds.) *The New American Political System*. AEI.
- Dowding, K. (1995). Model or metaphor: a critical review of the policy network approach. *Political Studies*, 43, 136-158.
- Hugh, P. (2000). Policy Network and Policy Learning: UK Economic Policy in the 1960s and 1970s. *Public Administration*, 78 (4), 771-792.
- Knoke, D., Pappi, F., and Tsujinaka, Y. (1996). *Comparing Policy Networks: Labor Politics in the U.S., Germany, and Japan*. Cambridge University Press.
- Marsh, D. (Eds.) (1998). *Comparing Policy Networks*. Open University Press
- OECD (2001). *Citizens as Partners: INFORMATION, CONSULTATION AND PUBLIC PARTICIPATION IN POLICY-MAKING*.
- Rhodes, R. and Marsh, D. (1992). New Direction in the Study of Policy Networks. *European Journal of Political Research*, 21, 181-205.
- Wasserman, S. and Faust, K (1994). *Social Network Analysis: Methods and Applications*. Cambridge University Press.
- Warden, F. (1992). Dimensions and Types of Policy Networks. *European Journal of Political Research*, 21, 29-52.

(2006.1.30受理)

The Evaluation and Analysis of a Policy about Information Flowing – The Case of Eco-Mark Program –

Takashi Hirahara

Kyoto University Graduate School of Global Environmental Studies
takashi.hirahara@k03.mbox.media.kyoto-u.ac.jp

Abstract

As design and the execution of a policy by a few policy actors became difficult, and the policy problems that they have deal with became complicated, various policy actors participated in a policy process, and the collection and disclosure of policy-related information became do. The evaluation of information approach policies and citizen's participation policy cannot enough analysis by existing evaluation approaches. Therefore I devised that I evaluated an information approach policies and citizen's participation policy in the basis by the policy network analysis that integrated quantitative society network analysis in mathematics sociology with a qualitative policy network theory in political science. In this thesis, Eco-labeling for use in informational approach is analyzed by policy network analysis method. Consequently, in the process of these analyses, it became clear to be different in a role of the policy actors to achieve for the information flowing by each eco-labeling goods, and between each structure of a network of the information flowing.

Keywords

informational approach, Social Network Analysis (SNA),
policy networks, eco-labeling, policy evaluation

【研究論文】

地方自治体における評価の波及と生成過程の分析

古川 俊一

筑波大学

furukawa@sk.tsukuba.ac.jp

森川 はるみ

茨城県庁

morikawa@ipu.ac.jp

要 約

本稿は、評価が自治体から自治体へと波及し、生成、実施されていく過程を分析することを目的とする。先行研究の検討から、波及・生成過程への寄与要因に注目した評価研究が新しいものであることを示し、全都道府県を対象としたイベント・ヒストリー分析と、10自治体を対象とした事例研究の結果をもとに、3つの知見を論じる。第1は、政治要因として、首長の熱意、行政改革の伝統及び政権交代の寄与が大きいこと、第2は、社会経済要因の寄与度は小さく、第3は、準拠集団という枠内での波及現象はみられない代わりに、全国規模での横並び競争が強力に機能していることである。総括すると、評価の波及と生成過程においては、全国的な横並びに加え、政治要因が重要であることが明らかになった。

キーワード

地方自治体、評価、業績測定、政策波及、イベント・ヒストリー分析、事例研究

序

本稿は、評価活動が自治体において波及・生成する過程を分析することを主な目的としつつ、分析を通じ得られた知見から、自治体が評価を導入・活用するに当たっての留意点を示唆する。

ここで、評価（evaluation）とは「プログラム又は政策の向上に寄与する手段として、その執行又は成果を、明示された又は暗黙の標準に照らして、体系的に評価（assess）すること」（Weiss 1998, p.4）としており、現在評価と呼ばれているものは、業績測定（performance measurement）も含めて、ほとんどこの定義に含まれる。本稿では、「行政評価」「政策評価」「事務事業評価」等様々な名

称で呼ばれ、実際に行われているものを幅広くとらえ、「評価」という用語で総称する（古川・北大路2004, p.327）。

1996年の三重県での評価導入以来、評価を導入する自治体が急速に増えてきた。総務省によれば、1999年度に評価を導入済であった自治体は、都道府県、政令指定都市、市町村の順に各々34%、8%、2%であったのが、2004年7月末の時点では、都道府県97.9%、政令指定都市100%、市区45.9%、町村6.7%へ増加している（総務省2004年調）。その背景には、地方分権の推進により、自己決定・自己責任を原則とする組織体制の整備が必要になったことや、厳しい財政事情の下での行政運営が求められていること等の理由により、行政運営の

成果が重視され始めたという事情がある（総務省2003、p.3）。

一般に、自治体が法的に義務づけられていない裁量的活動に取り組む場合、先行事例を参考にしながら開始することが多い。いわゆる「横並び」「前例」主義である。その場合、評価の目的を含めて、単純に模倣する傾向は否めない。評価の波及と生成過程を客観的に分析して両過程に寄与する要因を明らかにすることを通じて、今後の取り組みへの示唆が得られるだろう。

情報公開条例等に見られるように、まだこの自治体も採用していない新しい政策が他の自治体に順次採用され、全国的に広がっていく現象を政治学では政策波及（policy diffusion）と呼ぶ（伊藤2002、p.37）。本稿では、統計的な研究方法と事例研究の双方を用いる。

先行研究の検討から、波及・生成過程に寄与する要因に注目した評価研究が新しいものであることを示し、仮説を提示する。次にイベント・ヒストリー分析と事例研究の結果をもとに、知見を論じる。

1. 仮説の設定

(1) 先行研究の検討

戦後、昭和40年代までの評価の歴史には、大きく2つの流れがある。1つは公共事業の経済効果の測定であり、もう1つは行政効果の測定である（古川・北大路2004、p.47）。このうち行政効果の測定については、1953年に東京都で初めて取り組みがなされたが（黒沼1961）、組織の仕事全体に網をかけて広く浅く評価するという業績測定を組織的、全体的に実施したのは1996年以降の三重県が初めてである（古川・北大路2004、p.55）。その後、評価を導入する自治体が急速に増えたことを反映して、評価に関する出版が盛んになった。

日本の評価についての文献には、海外の事例を紹介したもの（上山1998; 1999） 予算、会計との関係から論じたもの（山本2001; 2002; 井堀2002） 費用便益分析や費用効果分析の観点から論じたもの（肥田野1997; 森杉1997） 評価マ

ニユアルとしての役割を意図したもの（山田2000; 小野・田淵2001） 個別自治体での事例を紹介したもの（村尾・森脇1999; 梅田2000; 2001; 2002; 上山・伊関2003） 日本の評価の現状をレビューしたもの（上山2002; 山谷2002） 日本における評価理論の体系化を図ったもの（山谷1997; 古川・北大路2004）がある。

評価の波及、生成過程についていえば、先行例を見ながら評価を導入する自治体が増加している実態を通じて波及過程が、また、三重県等での経緯を詳細に紹介した文献（梅田2000; 2001; 2002; 上山・伊関2003）や評価の成功条件に触れている文献（山谷1997; 古川・北大路2004）を通じて生成過程が明らかになりつつあるという解釈も成り立つ。しかし、両過程に寄与する要因は何かという観点から、精緻な分析を行った例はまだない。

両過程に寄与する要因を探る上で、一般的に自治体の政策決定要因は、政治要因、社会経済要因、波及要因の3つに分類できる（例えば、Gray 1994）。前2者は、自治体固有の属性に由来する内的要因であり、波及要因は、他自治体の動向という外部条件に由来する外的要因である。この3つの要因が情報公開等の政策波及に及ぼす影響を分析した研究は日本でもなされているが（伊藤2002） 評価を対象にした研究例はまだない。

(2) 仮説

政治要因、社会経済要因及び波及要因ごとに仮説を提示する。仮説の提示に当たっては、(1)で参照した個別事例の記述や成功条件を分析した文献も参照している。

まず、第1の政治要因については、次の8つを挙げる。仮説の設定に際しては、伊藤（2002）を参考にした。

1番目は首長の党派性である¹。高度成長期において新政策を自治体が先導した例を首長の革新性と結びつけた議論があるが（中西1979; Steiner 1980） 近年は、選挙時にどの政党の支援も受けないいわゆる「無党派」が広がる傾向にあり、党派性の影響力が弱まっていることから、首長の党派性は評価の導入に影響しないと予想できる。

2番目は首長の経歴である。片岡（1994）は、

首長が中央官僚出身であると新政策の採用が早い可能性を示唆している。確かに中央とのつながりがあれば、政策アイデアが中央から即時的に伝わるのが可能である。伊藤（2002）が指摘したように、国の介入によって政策波及が早まることもあるため、評価についても、国の法制化の動きを早い段階で知ることと普及速度が速まった可能性がある。

3番目は首長の当選回数である。梁（1991）は当選回数が増えるほど市長の影響力が高まるとしているが、伊藤（2002、p.70）は多選による影響力と現状維持の傾向が打ち消し合うと主張している。評価の場合、自らの行政運営の成果を見直すという考え方は、現状維持の傾向が強まるほど受け入れ難くなると予想されることから、むしろ、当選回数が少ない段階の方が評価に取り組みやすい可能性がある。

4番目は評価にかける首長の熱意である。自治体における影響力構造の研究では、首長が政策決定に与える影響力が高く評価されており（中野1986; 小林他1987）首長が自ら熱心に評価への取り組み姿勢を示すことが評価への取り組みを推進すると予想される。

5番目は首長の選挙年か否かである。例えば、公約として評価の導入を掲げて当選する場合、選挙年であることが評価導入の強い動機付けとなることから、選挙年直後は評価の導入が進むと考えられる。

6番目は議会の影響である。議会の影響は重要であり、与党が多数を占めるか否かで政策形成の結果が変化する（小林他1987; 中野1992）。しかし、評価は条例という形を取らない限り議会審議を必須としないこと、公共事業と異なり、議員と直接の利害関係にないことから、議会の勢力比によって評価の導入が左右されることはないかと予想される。

7番目は行政改革の伝統である。自治体内部に積極的に行政効率化を求める動きや新しいものを受け入れる土壌があれば、評価の重要性を素早く理解できることから、評価の受容が速やかに進むと予想される。

8番目は政権交代である。三重県のように、旧政権の継承者を破って当選した新しい首長が自ら

の新規性を打ち出すための手段として評価を活用している例があることから、政権交代が生じた場合、評価への取り組みが一気に進むことがありうる。

したがって、政治要因についての仮説は【政1】～【政8】のとおりである。

【政1】首長の党派性は評価の導入に影響しない

【政2】首長が中央出身であると評価を導入する

【政3】首長の当選回数が少ないほど評価を導入する

【政4】首長の熱意があると評価を推進する

【政5】選挙年直後に評価を導入する

【政6】議会の構成は評価の導入に影響しない

【政7】行政改革の伝統があると評価を推進する

【政8】政権交代の結果として評価を推進する

第2の社会経済要因については、自治体規模（Knoke 1982）、都市集積度（飽戸・佐藤1986）、財政的余剰（Reed 1986）が大きいと政策決定が促進されるとする研究を参考にしながら、高度経済成長時代の財政と異なり財政の逼迫が緊急であることから、財政の効率性の状況も要因としてありうると考え、次のように設定する。

【社1】自治体規模が大きいほど評価を導入する

【社2】都市化が進んだ自治体ほど評価を導入する

【社3】地域が豊かなほど評価を導入する

【社4】財政の効率性が低いほど評価を導入する

第3の波及要因では、政策の相互参照を行う自治体集団である準拠集団内においても、情報交換を定期的に行い政策形成の参考としている実態を反映して、横並び競争が生じていると考える（村松1986; 1988; 伊藤2002）。

【波1】全国導入数が多いほど評価を導入する

【波2】準拠集団内の導入数が多いほど評価を導入する

2. 研究のデザイン

(1) 分析手法

評価の波及過程は、三重県に始まった評価が次々と他の自治体に導入されていくプロセスであり、プロセスの性質上、特定の自治体だけに注目

しても全体的な傾向がつかめないことから、一定レベルの自治体全てを対象に統計的手法を用いて全体的な傾向を把握することが適切である。

一方、評価の生成過程は、各自治体で評価の導入が決定され、実施されていくプロセスであり、プロセスの性質上、個別の自治体の事例を詳細に分析する手法が適している。

したがって、本稿では、統計的な研究手法と事例を個別に見ていく研究手法の双方を用いることとし、イベント・ヒストリー分析と事例研究の双方の手法を用いた。政治要因、社会経済要因、波及要因のそれぞれの寄与度に関する仮説実証型の研究である。

評価導入時期については、本格的な実施の前に試行期間を設けている場合や、北海道の「時のアセスメント」が「政策アセスメント」へ進化したように、一旦評価を導入しても、数年後に評価の形態を変化させている場合がある。したがって、どの時点で評価が導入されたといえるのか、判断の基準を設ける必要がある。情報公開や行政手続については、条例が成立した時期をもって導入とすることもできるが（伊藤2002）評価の場合、都道府県で条例を制定しているのは、2004年7月末の時点で北海道、岩手県、宮城県、秋田県の4県のみであり（総務省調）まだ一般的ではない。したがって、本稿では、試行段階を経て、評価が全庁的に本格的に実行された時点を導入時期と考え、分析対象とした。

(2) イベント・ヒストリー分析

政策波及研究は、米国州政治についての研究から発展した（Walker 1966; Gray 1973）。日本においても、村松（1975; 1988）が、政治に後押しされた横並び競争や革新自治体が政策波及に果たす役割を明らかにした研究を行っている。

他方、イベント・ヒストリー分析（Event History Analysis、以下「EHA」）は、事象（event）が生起するパターンを分析する手法であり（Yamaguchi 1991, p.1）新政策がある時点で採用されたかどうかを従属変数とし、それを採用者の属性、時間、環境等に関する独立変数で説明していく統計分析手法である。EHAは時間的、空間

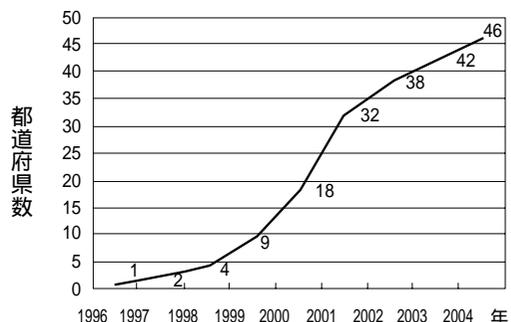
的に広がりのある波及現象の分析に適している（伊藤2002, p.257）。

EHAは、元々は疫学における生存分析から発展し、社会学（Tolbert and Zucker 1983; Rogers 1995）政治学・行政学（Berry and Berry 1990; 1992; Mintrom 1997）に応用されてきている。日本の地方自治体研究に応用した例には、伊藤（1999; 2002）がある。ここで地方自治体における評価について考えると、行政分野をまたいで横断的に適用されるとはいえ、評価の導入はそれ自体が政策と考えられるため、先行研究を踏まえ、政策波及研究の枠組を適用することが可能である。また、都道府県の評価は、図1のようにS字曲線の形で普及が進んでおり、普及パターンの典型例（Rogers 1995, p.356）であるから、EHAの活用にも適している。

EHAで重要な概念はハザードレート（hazard rate: 危険率）であり、これは、ある時点で事象が生起する瞬間的な確率を意味する（Yamaguchi 1991, p.3）。ハザードレートが時間の経過と共にどのように変化すると捉えるかによって、Weibullモデル等複数のモデルが存在する（Blossfeld and Rohwer 2001）。

EHAでは、事象が生起するまでの期間には、事象が生じ得ない期間（period not at risk for having the event）と事象が生じうる期間（period at risk）の2つがある（Yamaguchi 1991, p.2）。本稿では、評価をまだ導入していない全ての自治体において導入の可能性があるため、まだどの自治

図1 都道府県における評価累計導入数



（出所）各都道府県ホームページ（2005年8月1日現在）
（注）本格導入時期で集計（試行導入は除く）

体も導入を開始していない1995年を分析の起点とする。

(3) 事例研究

事例研究は、幅広い記述や分析を通じて事例を深く包括的に理解する方法である。事例研究の一般化可能性を高めるためには、構成概念妥当性、内的妥当性、外的妥当性、信頼性の4つの基準を満たす必要がある (Singleton et al. 1988; Yin 1994)。本稿では、一般化可能性を高めるために、インタビューや文献等複数の情報源を活用し、どの情報源からどの情報を得たかを明らかにする。また、複数の事例を通じて傾向を把握する等、個々の事例の記述にのみ焦点を当てないよう留意する。

(4) 分析の対象

EHAでは都道府県のみを対象とする。近年、市区での評価導入が急速に進んでいるが、未だ45.9% (総務省2004年調) と半数程度の導入であること、データ収集のコストがかかること及び同レベルの自治体を対象とすることが望ましいことが理由である。

一方、事例研究については、北海道、宮城県、秋田県、神奈川県、静岡県、三重県、草加市、杉並区、三鷹市、横須賀市の10自治体を対象とする。これは、事例選択の3つの基準 (GAO 1990, p.22)、すなわち利便性 (convenience)、目的合致性 (purpose) 及び事例選択の確率 (probability) による。都道府県と市町村の両方を含むこと、評価先進自治体である静岡県、三重県、草加市、三鷹市を含むこと、事務事業評価の代表である三重県、業務棚卸法の代表である静岡県、外部評価の代表である三鷹市等、評価の内容が様々な自治体を含むこととした。EHAと異なり、事例研究では都道府県と市町村の双方を扱い、より一般的な考察を志向する。

3. イベント・ヒストリー分析

(1) 分析方法

ア 定式化

政治要因に関する変数を $x(p_i)$, $i=1, \dots, 5$ 、社会経済要因に関する変数を $x(s_j)$, $j=1, \dots, 8$ 、波及要因に関する変数を $x(d_k)$, $k=1, 2$ と表現する。評価導入の有無を y とすると、定式は、

$$y = \sum_{i=1}^5 x(p_i) + \sum_{j=1}^8 x(s_j) + \sum_{k=1}^2 x(d_k)$$

となる。

これは、評価導入の有無が表1の変数によって決定されることを表したものである。

イ 変数設定

変数の詳細及びデータは、表1のとおりである。ただし、首長の熱意、行政改革の伝統及び政権交代に関する仮説【政4】【政7】【政8】はデータの設定が困難であるため、変数設定を行わずに事例研究で対応する。

ウ モデル

本稿で扱うデータは、年単位という離散データであり、かつ、データ収集の都合上、2002年で打ち切られた打ち切りデータ (censored data) である。離散データの場合、離散時間Logitモデル (discrete-time Logit model) が比例ハザードモデル (proportional hazards model) を適用する主張があり (Yamaguchi 1991, p.15)、離散データを用いた伊藤 (2002) もLogitモデルを使用している。ただし、伊藤 (2002) はハザードレートが年と共に変化すると考えて時間ダミー変数を加えているが、変数「全国導入数」と時間ダミー変数を共にLogitモデルに加えると、多重共線性が生じる可能性が指摘されている (Beck, Katz and Tucker 1998, p.1273)。²⁾

一方、多重共線性を回避するため時間ダミー変数を削除すると、2002年のデータ打ち切り時点でまだ評価を導入していなかった自治体のその後の導入確率についての情報を分析に加えることがで

表1 イベント・ヒストリー分析の変数

変数		内容
政治要因: $x(p_i), i=1, \dots, 5$		
首長の党派性	$x(p_1)$	保守系首長 革新系首長 相乗り首長 (ダミー変数)
首長の経歴 (注1)	$x(p_2)$	中央官僚、地方政治家等(ダミー変数)
首長の選挙時期	$x(p_3)$	選挙年 選挙前年 選挙翌年(ダミー変数)
議会の政党勢力比	$x(p_4)$	自民党議席割合(%)
	$x(p_5)$	拮抗議会(ダミー変数)
社会経済要因: $x(s_j), j=1, \dots, 8$ (注4)		
規模	$x(s_1)$	人口(万人)
都市化度	$x(s_2)$	DID(注2)人口比(%)
都市化速度	$x(s_3)$	人口増減率(%)
産業集中	$x(s_4)$	1人当たり製造品出荷額(百万円)
豊かさ	$x(s_5)$	1人当たり県民所得(千円)
財政の効率性と自律性	$x(s_6)$	1人当たり歳出額(百万円)
	$x(s_7)$	財源依存度(%)
	$x(s_8)$	財源依存度変化率(%)
波及要因: $x(d_k), k=1, 2$ (注4)		
政策波及	$x(d_1)$	全国導入数
	$x(d_2)$	準拠導入数

(出所)伊藤(2002)をもとに筆者作成。ただし、「財源依存度変化率」を追加。
 (注1)中央官僚、自治官僚、建設農水官僚、地方官僚、中央政治家、地方政治家、知事当選回数7つの属性が含まれる。それぞれを1(経験有り)・0(経験なし)とコード化して分析に用いた。知事当選回数のみは、当選回数そのものをコード化した。
 (注2)人口集中地区(Densely Inhabited Districts)のこと。国勢調査で用いられている概念
 (注3)歳出総額は、普通会計(一般会計と公営事業会計を除く特別会計との合計)額である。
 (注4)社会経済要因は前年のデータであり、波及要因である全国導入数と準拠集団ごとの導入数は、前年までの累計導入数

(注5)地域ブロックを基本とするが、政令指定都市を有する10道府県に東京都と埼玉県を加えた12都道府県は「大都市集団」として別に扱う。大都市集団に属する自治体の場合、準拠集団は地域ブロックではなく、「大都市集団」となる。これは伊藤(2002)に基づいている。

きない。

したがって、本稿では、打ち切りデータを扱うのに適したWeibullモデルを用いることで、より精密な分析を行う。Weibullモデルとは、スエーデンのW.Weibullが鋼球の寿命を解析するために開発したWeibull分布を仮定したモデルである。信頼性工学の分野では破壊現象の解析に有効とされており、本稿のような右側打ち切りデータの解析に適している。

また、比例ハザードモデルについては、本稿のような年単位データの場合、評価導入という事象が頻繁に同時発生するためにバイアスが生じる可能性が高いことから(Yamaguchi 1991, p.16)本稿では用いない。

なお、社会経済要因については、多重共線性が発生する可能性があることを確認した。したがって、社会経済要因に含まれる8つの変数を互いに排除した8つの定式について分析を行う。具体的には以下のとおりである。j=1の場合を定式(1)とし、以下、順に定式(2)~(8)とする。

$$y_j = \sum_{i=1}^5 x(p_i) + x(s_j) + \sum_{k=1}^2 x(d_k) \quad (j=1, \dots, 8)$$

(2) 分析結果

EHAの結果を表2に示す。10%水準で有意な結果のみを記述した。なお、統計ソフトはS-plus2000を用いている。

10%水準で見ると、全ての定式で「全国導入数」が有意となり、定式(3)では他に「中央政治家」が、定式(6)では他に「中央政治家」と「1人当たり歳出額」が有意となった。「中央政治家」は、首長の経歴に中央政界での政治経験が含まれることを意味する。「1人当たり歳出額」は、財政の非効率性を表す変数である。

表2 イベント・ヒストリー分析の結果

定式	(1)	(2)	(3)
政治要因			
中央政治家			-0.096* (0.058)
波及要因			
全国導入数	0.014*** (0.003)	0.013*** (0.003)	0.014*** (0.003)
適合度	48.96***	50.14***	50.09***
定式	(4)	(5)	(6)
政治要因			
中央政治家			-0.106* (0.058)
社会経済要因			
1人当たり歳出額	-	-	0.343* (0.177)
波及要因			
全国導入数	0.014*** (0.003)	0.014*** (0.003)	0.012*** (0.003)
適合度	49.45***	49.41***	53.33***
定式	(7)	(8)	
波及要因			
全国導入数	0.014*** (0.003)	0.015*** (0.003)	
適合度	48.91***	49.15***	

(出所)筆者作成。
 (注1) *:p<0.1 **:p<0.05 ***:p<0.01
 (注2) 値は係数、括弧内は標準誤差
 (注3) 適合度はカイ二乗適合度検定の数値
 (注4) 小数点以下第4位四捨五入

波及要因である「全国導入数」が非常に強力に作用している。情報公開条例、行政手続条例等の政策波及と研究にEHAを用いた伊藤(2002)においても、同じ結果が示されている。このような場合、「全国導入数」の働きによって他の変数の効果が限定されている可能性を否定できない。そのため、「全国導入数」と他の変数間での交互効果(interaction effect)を見ることにした(表3)。

その結果、新たに「知事当選回数」と「選挙年」の2つの変数が有意となったが、「全国導入数×知事当選回数」と「全国導入数×選挙年」の交互効果が10%水準で有意となっているため、変数の主効果は交互効果によって限定される。

解釈を深めるため、交互効果を図示するための操作を行った。手順は次のとおりである。まず、変数「知事当選回数」「選挙年」「全国導入数」の平均と標準偏差を求め、それを元に、各変数の値

表3 交互効果の結果

	知事当選回数	選挙年
政治要因		
中央政治家	-0.093* (0.056)	-0.103* (0.058)
知事当選回数	0.055***(注4) (0.025)	
選挙年		0.149*(注5) (0.087)
社会経済要因		
1人当たり歳出額	0.334*(注4) (0.172)	0.318* (0.172)
波及要因		
全国導入数	0.020*** (0.005)	0.015*** (0.003)
交差項		
全国導入数×知事当選回数	-0.003** (0.001)	-
全国導入数×選挙年	-	-0.010* (0.005)

(出所)筆者作成
 (注1) *:p<0.1 **:p<0.05 ***:p<0.01
 (注2) 値は係数、括弧内は標準誤差
 (注3) 小数点以下第4位四捨五入
 (注4) 全国導入数との交互効果における知事当選回数に付与された係数
 (注5) 全国導入数との交互効果における選挙年に付与された係数

表4 「知事当選回数」「選挙年」「全国導入数」の平均値と標準偏差

変数名	平均値	標準偏差
知事当選回数	2.483	1.425
選挙年	0.240	0.428
全国導入数	6.281	8.374

(出所)筆者作成
 (注) 小数点以下第4位四捨五入

表5 「知事当選回数」「選挙年」「全国導入数」の値の変動

変数名	High (平均値+標準偏差)	Low (平均値-標準偏差)
知事当選回数	3.907	1.058
選挙年	0.667	-0.188
全国導入数	14.655	-2.092

(出所)筆者作成
 (注1) 小数点以下第4位四捨五入
 (注2) 各変数の値の分布を示しているため、負の値をとることもあり得る。

の変動をHighとLowで表す。結果は表4及び表5のとおりである。

次に、各変数がHighとLowの値を取る場合の推定値（評価導入の可能性）を求める。表3の係数と表5の値を $Y = (\text{定数項}) + (\text{変数1}) + (\text{変数2}) + (\text{変数1} \times \text{変数2})$ の推定式に当てはめることによって、「知事当選回数」と「全国導入数」の値の変動が評価導入の可能性に及ぼす影響を数値で示すことができる。変数1には「知事当選回数」または「選挙年」を当てはめ、変数2には「全国導入数」を当てはめる。

推定の結果は表6及び表7のとおりである。数値が大きいほど、評価導入の可能性が高いことを示す。

表6 「知事当選回数」「全国導入数」の値の変動による評価導入可能性

		全国導入数	
		Low	High
知事当選回数	Low	1.242	1.538
	High	1.416	1.579

(出所) 筆者作成
(注) 小数点以下第4位四捨五入

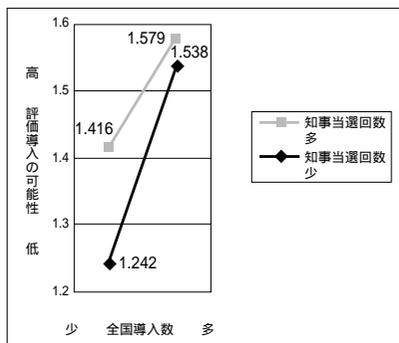
表7 「選挙年」「全国導入数」の値の変動による評価導入可能性

		全国導入数	
		Low	High
選挙年	Low	1.234	1.519
	High	1.379	1.523

(出所) 筆者作成
(注) 小数点以下第4位四捨五入

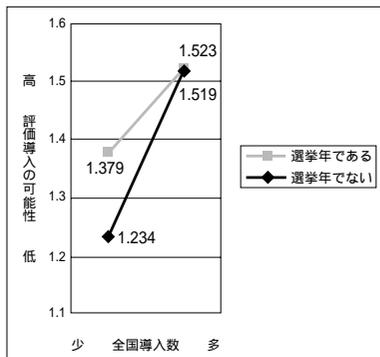
以上の結果を図示すると、図2及び図3のようになる。

図2 「知事当選回数」と「全国導入数」の交互効果



(出所) 筆者作成

図3 「選挙年」と「全国導入数」の交互効果



(出所) 筆者作成

以上の結果を解釈すると、評価の普及が進み、全国での導入数が多くなった時期には、知事当選回数や選挙年は、評価の導入にあまり影響しない。しかし、全国導入数がまだ少ない時期には、評価の導入を促進する傾向にあるということが出来る。

(3) 考察

第1のカテゴリーである政治要因については、首長の党派性と議会の構成は評価の導入に影響を及ぼさないことが示された。首長の経歴では、中央政治家であった経歴が負の方向に働いた。これは、むしろ中央とのつながりのない首長の方が思い切った政策を打ち出すことができることを表しており、最近の地方分権の推進やいわゆる無党派知事の登場と呼応するものである。

また、交互効果の分析により、次の2点が明らかになった。第1点は、全国導入数が少ない場合、当選回数を重ねて政権基盤を強化した首長の方が評価を導入しやすい。第2点は、全国導入数が少ない場合、知事の選挙年であると評価を導入する可能性が高い。ただし、多数の自治体での導入が進んでくると、当選回数や選挙年に影響されることがなくなる。

第2に、社会経済要因については、財政の効率性が低いほど評価を導入するという仮説が一部支持された。財政の非効率性を表す変数である「1人当たり歳出額」が多いほど、評価を導入する可能性が高い。だが、その他の社会経済要因で有意

となる変数が得られなかったことは、評価の場合、自治体規模、都市化度、地域の豊かさ等の指標が導入の決定に影響しないことを意味する。どの自治体が先行して政策を導入するかは、社会経済的な指標だけでは判断できないということである。

第3に、波及要因は、前年までに全国でどれだけの自治体が評価を導入したかという情報が強力に影響することが示された。一方、準拠集団の情報は影響しないようである。これは、情報伝達が質量共に活発化している現状においては、準拠集団という枠にとらわれず、広く全国規模での横並びが生じている状況を示している。

なお、このEHAの分析結果は都道府県のみを対象としたものであり、市町村への適用には留意する必要がある。

4. 事例研究

(1) 調査方法

2003年8月15日～28日の期間に対象自治体を訪問し、評価担当幹部又は担当者に1時間程度のインタビューを実施した。また、これ以外の期間において、知事または市長もしくはそれに準ずる幹部その他の関係者に対するインタビュー、文献調査等、複数の情報源により調査を実施した。

対象自治体インタビューにおける質問項目は次の2つである。「評価導入時～現在に至るまで、政治的環境（首長のご意向、議会との関係など）が、評価のあり方にどのように影響を及ぼしてきましたか。」「現時点での、評価の効果／問題点／課題についてお聞かせください。」ただし、結果として聴取できた内容はこの2点を超えるものが多い。

知事または市長もしくはそれに準ずる幹部その他の関係者に対するインタビューは、あえて事前の質問項目を設けず、フリートーク形式で評価に関する多様な面からの聞き取りを行った。

(2) 分析結果

自治体ごとに、「評価のレベル」「行政改革の伝統」「政権基盤」「議会の関与」「首長の熱意」「職

員の熱意」の6つの項目についてレベル付けを行った（表8）。レベル付けの対象時期は評価導入時とし、評価のレベルが高い順に並べ替えている（したがって、「政権基盤」以外は、年度経過にしたがって微妙にレベルが変わる可能性があることに注意）。各項目のレベル付けは、表9の基準に従って行った。

このうち、評価のレベルは、公共経営改革の結果を業務結果、管理又は意思決定の過程の改善、政治行政システムの能力の変化、システム全体の移行、の4つに分類したPollitt and Bouckaert (2004, p.98)を参考にしている。レベル1は、人員、業務量、コスト等の削減が見られるなど、数値で表せる範囲での行政活動の効率化を意味する。レベル2はさらに一歩進んで、行政組織内での意思決定過程がより効率的なものに変容する水準までも指す。レベル3は、改革派の首長が就任する等を契機に、末端まで改革精神が浸透していくような段階である。

レベルの数値が高いほど、その項目が強力に出現していることを意味する。

表8 事例研究の分析結果

	評価のレベル	行政改革の伝統	政権基盤		議会の関与	首長の熱意	職員の熱意	
			政権交代	政権継承			職員の取り組み	キーパーソンの存在
三鷹市	3	2	×		1	2	2	
静岡県	3	2	×		1	2	2	×
横須賀市	3	1	×		1	3	2	
三重県	3	0		×	0	3	2	
草加市	2	0	×		1	2	2	
秋田県	1	0		×	0	2	1	×
宮城県	1	0		×	0	2	1	×
杉並区	1	0		×	1	2	1	×
神奈川県	1	1	×		1	1	0	×
北海道	1	0	×		0	2	1	×

（出所）筆者作成。

（注）数値はレベルを表す。

表9 レベル付けの基準

項目	レベル0	レベル1	レベル2	レベル3
評価のレベル	何も変化がみられない場合	人員、業務量等に变化がみられた場合	意思決定過程が改善された場合	行政文化に変容がみられた場合
行政改革の伝統	行政改革の伝統がみられない場合	従来から先進的な取り組みを行ってきた場合	従来から行政改革の伝統が存在する場合	-
政権交代	評価生成時の首長が政権交代によってトップの座に就いた場合は、そうでない場合は×			
政権継承	評価生成時の首長が比較的安定的な政権継承によってトップの座に就いた場合は、そうでない場合は×			
議会の関与	議会が評価に関心を示さない場合	議会が評価の実施を要望する場合	議会が評価を実施するよう圧力をかける場合	-
首長の熱意	首長が評価に関心を示さない場合	首長が評価の導入を指示した場合	首長が積極的に評価に組み込む意欲を示した場合	首長が評価の導入によって組織文化の向上をも意図した場合
職員注の取り組み	職員が評価に関心を示さない場合	一部の職員が評価の意義を理解している場合	多数の職員が評価の意義を理解している場合	-
キーパーソンの存在	職員にキーパーソンが存在する場合は、存在しない場合は×			

(出所) 筆者作成。

(注) ここでの職員とは、役職を問わず、広く職員一般を指す。

この結果から、次の4つのことが分かる。第1に、「行政改革の伝統」は、評価の生成に非常に強力に働いているということである。第2に、政権基盤については、「政権交代」と「政権継承」のどちらがより有効に作用するかを判断するのは難しいということである。例えば、三重県と宮城県では、同じ政権交代という状況であっても、評価のレベルが異なって現れている。第3に、議会の影響が定かでないということである。評価は国の行政改革の流れに合致するものであり、世論にも評価を歓迎する気風があることから、議会の構成や首長との関係にかかわらず、賛意を表明する機会が多いと解釈できる。第4に、「首長の熱意」と「職員の熱意」の両者が揃った場合に、評価が高いレベルを示すということである。

(3) 考察

【政4】首長の熱意があると評価を推進するという仮説は一部支持された。例えば、宮城県、杉並区では、首長の熱意はレベル2であっても職員の熱意は必ずしも高くなく、評価のレベルは低い。一方、三重県と横須賀市では首長の熱意と職員の熱意の双方が強力で、評価のレベルも高いという結果であった。三重県と横須賀市では、首長の熱意がレベル3で、行政全体のマネジメントを改革しようという首長の強い意思が働いており、首長のマネジメントへの関心が評価を高レベルとした可能性がある。だが、宮城県、杉並区等の結果をみると、首長の熱意だけでは不十分であり、首長の考えが職員に浸透し、職員自ら評価の重要性を理解するに至らないと、評価は効果を十分に発揮することができないと解釈できる。

次に、【政7】行政改革の伝統があると評価を推進するという仮説は支持された。静岡県と三鷹市の場合、評価の導入自体は特段新しいものではなかった。なぜなら、従来から絶え間ない行政改善の試みが続けられてきていたため、評価についても改善の一環と受け取られ、行政活動の成果を指標によって把握して効率的な行政運営に役立てるという趣旨が円滑に首長と庁内に浸透したからである。静岡県も三鷹市も、三重県のような強烈的な首長のリーダーシップはうかがえず、静かな改革

の流れであるが、元々の改革意識が評価の生成にいかにか重要であるかが示された。改革意識を根付かせることが評価生成の第一歩である。

最後に、【政8】政権交代の結果として評価を推進するという仮説は一部支持された。杉並区では公約どおり評価を導入したが、評価によって政権基盤を強化しようという意思は読み取れなかった。一方、三重県、宮城県、秋田県の場合は、政権交代という形で初当選し、かつ、当初は政権基盤が弱かったことから、評価を自らの基盤確立のために活用する意思が働いた。ただし、この3県の中で評価が高レベルなのは三重県だけであることから、政権交代は評価導入のきっかけとなるが、評価のレベルを高めるには、首長が評価の導入によって組織文化の向上をも意図するほど熱心であること、さらに、首長の熱意と職員の熱意の双方が揃うことが必要だと解釈できる。この双方が揃った例として、導入時のレベルはさほど高くなかったが、数年を経てキーパーソンが現われはじめ政策形成のパターンを変化させ、条例化を行うなど制度化を進め、レベル2に近づいていったとみられる宮城県、秋田県がある。

結論

イベント・ヒストリー分析と事例研究の結果を統合すると、表10のとおりとなる。

以上の分析結果から、3つの結論を導いた。

第1は、政治要因として、首長の熱意、行政改革の伝統及び政権交代が評価の波及と生成過程に寄与していることである。両過程は自治体固有の政治要因に起因して進行してきたことになる。中でも、行政改革の伝統という要因が寄与していることは、自治体が行政運営の改善に積極的に取り組む姿勢が重要であることを示している。

第2は、社会経済要因が評価の波及と生成過程に及ぼす寄与度の小ささである。Dye and Robey (1980) や飽戸・佐藤 (1986) によって、政策決定過程に及ぼす社会経済要因の重要性が示されてきたが、地方自治体の評価については当てはまらないことになる。その理由としては、大都市圏での急激な財政悪化による地域間経済格差の縮小等、近年の社会経済的な状況の変化が挙げられる。

表10 仮説の検証結果

仮説		結果
政治要因		
政 1	首長の党派性は評価の導入に影響しない	支持
政 2	首長が中央出身であると評価を導入する	不支持
政 3	首長の当選回数が少ないほど評価を導入する	不支持
政 4	首長の熱意があると評価を推進する	一部支持
政 5	選挙年直後に評価を導入する	一部支持 (全国導入数が少ない時)
政 6	議会の構成は評価の導入に影響しない	支持
政 7	行政改革の伝統があると評価を推進する	支持
政 8	政権交代の結果として評価を推進する	一部支持
社会経済要因		
社 1	自治体規模が大きいほど評価を導入する	不支持
社 2	都市化が進んだ自治体ほど評価を導入する	不支持
社 3	地域が豊かなほど評価を導入する	不支持
社 4	財政の効率性が低いほど評価を導入する	一部支持
波及要因		
波 1	全国導入数が多いほど評価を導入する	強力に支持
波 2	準拠集団内の導入数が多いほど評価を導入する	不支持

(出所) 筆者作成

1996年以降全国で見られた評価の波及と生成過程は、財政の効率性が低いほど評価を導入する可能性が高いものの、全体的な傾向としては、予想に反し、社会経済要因によることなく進行してきたと解釈できる。

第3は、波及要因の寄与度が非常に大きいことである。ただし、評価の場合、情報伝達が質量共に活発化している現状を反映して、準拠集団という枠内での波及現象はみられなかった。分析結果として示されたのは、全国規模での横並び競争が強力に機能していることであり、村松 (1975) 以来示されてきた自治体の横並び競争が評価の波及

過程においても強力に寄与していることが分かった。

総括すると、評価の波及と生成過程において、全国的な横並びに加え、政治要因が重要であることが明らかとなった。特に、評価の有効な活用の度合いは、首長の熱意、行政改革の伝統及び政権交代に相当依存している。このことは何を意味するか。第1に、導入に当たっても展開に当たっても、トップダウン型の契機は必要不可欠である。およそ大方針はトップの理解がなければ進まない。第2に、しかし、トップの方針だけでは制度化されないし、骨抜きになって定着しない。中間管理職の貢献、改革に取り組む組織風土や伝統がなければ持続しないし、形式的な書類作成の弊に陥る。第3に、アクターが変わり、外的圧力が加わるという意味で、政権交代も1つの要因足りうるのである。

注記

- 1 今回の分析対象期間（1995～2002年）は、保守・革新の別が弱まり、選挙時にどの政党の支援も受けなないいわゆる「無党派」が広がり始めた時期に当たる。一般には、「無党派」をいわゆる「改革派」と見なす向きもあるが、その捉え方は曖昧である。本稿では、「党派性」を従来の保守・革新の流れを汲むものと定義し、時代とともに党派性の影響力が弱まっていると考え、仮説を提示した。
- 2 Masuyama (2002)は、変数「全国導入数」を加えることそのものが多重共線性を引き起こすと指摘している。しかしながら本稿では、その可能性を考慮しつつも「全国導入数」の重要性に鑑み、交互効果の検証を行い、慎重にモデルに組み込むことで対処した。なお、「全国導入数」を除いてEHAを行うと、他のほぼ全ての変数が有意となったが、これは、「全国導入数」という重要な変数を取り去ることの弊害を示していると解釈した。

謝辞

本稿の草稿について建設的な批評をいただいた伊藤修一郎筑波大学教授、及び詳細かつ有益なコメントを寄せてくださった匿名の査読者に感謝する。ありうべき誤りは、当然筆者の責任である。

参考文献

- 鮑戸弘・佐藤誠三郎（1986）「政治指標と財政支出647市の計量分析」、大森彌・佐藤誠三郎編『日本の地方政府』、東京大学出版会
- 伊藤修一郎（1999）「情報公開条例の制定過程」、『年報公共政策』、2号
- 伊藤修一郎（2002）『自治体政策過程の動態：政策イノベーションと波及』、慶応義塾大学出版会
- 井堀利宏（2002）「望ましいガバナンスへの移行戦略：予算システム」、宮川公男・山本清編『パブリック・ガバナンス：改革と戦略』、日本経済評論社、152-165
- 上山信一（1998）『「行政評価」の時代：経営と顧客の視点から』、NTT出版
- 上山信一（1999）『「行政経営」の時代：評価から実践へ』、NTT出版
- 上山信一（2002）『日本の行政評価：総括と展望』、第一法規出版
- 上山信一・伊関友伸（2003）『自治体再生戦略：行政評価と経営改革』、日本評論社
- 梅田次郎（2000）「政策評価導入の体験：事務事業評価システムの導入に対する三重県庁内の組織的抵抗」、『日本公共政策学会年報』CD-ROM版
- 梅田次郎（2001）「業績測定型・三重県事務事業評価システムの発展過程と展望」、『日本評価研究』、1（2）：69-77
- 梅田次郎（2002）「意識改革と政策形成：三重県庁における自治体組織運営の変革プロセス」、『公共政策研究』、有斐閣、55-69
- 小野達也・田淵雪子（2001）『行政評価ハンドブック』、東洋経済新報社
- 片岡正昭（1994）『知事職をめぐる官僚と政治家』、木鐸社
- 黒沼稔（1961）「行政効果の測定」、『自治研修』、25:37-51
- 小林良彰・新川達郎・佐々木信夫・桑原英明（1987）『アンケート調査にみる地方政府の現実：政策決定の主役たち』、学陽書房
- 総務省（2003）『地方公共団体における行政評価の導入の実態と今後の展開について：平成14年度地方公共団体における行政評価についての研究会報告』、3月

- 中西啓之 (1979) 「日本の革新自治体」, 島恭彦・池上惇・遠藤晃 『現代の地方自治』, 自治体研究社
- 中野実編 (1986) 『日本型政策決定の変容』, 東洋経済新報社
- 中野実 (1992) 『現代日本の政策過程』, 東京大学出版会
- 肥田野登 (1997) 『環境と社会資本の経済評価: ヘドニック・アプローチの理論と実際』, 勁草書房
- 古川俊一 (2004) 「地方行革の新展開と経営志向の台頭 NPMと評価の効用と限界」, 『行政管理研究』, 105: 45-56
- 古川俊一・北大路信郷 (2004) 『新版 公共部門評価の理論と実際』, 日本加除出版
- 村尾信尚・森脇俊雄 (1999) 『動き出した地方自治体改革』, 関西学院大学出版部
- 村松岐夫 (1975) 「戦後日本の地方政治」, 足立忠夫他編 『現代政治と地方自治』, 有信堂
- 村松岐夫 (1986) 「政府間関係と政治体制」, 大森彌・佐藤誠三郎編 『日本の地方政府』, 東京大学出版会
- 森杉壽芳編 (1997) 『社会資本整備の便益評価: 一般均衡理論によるアプローチ』, 勁草書房
- 山田治徳 (2000) 『政策評価の技法』, 日本評論社
- 山本清 (2001) 『政府会計の改革』, 中央経済社
- 山本清 (2002) 「望ましいガバナンスへの移行戦略: 会計システム」, 宮川公男・山本清編 『パブリック・ガバナンス: 改革と戦略』, 日本経済評論社, 166-179
- 山谷清志 (1997) 『政策評価の理論とその展開: 政府のアカウントビリティ』, 晃洋書房
- 山谷清志 (2002) 「わが国の政策評価: 1996年から2002年までのレビュー」, 『日本評価研究』, 2(2):3-15
- 梁起豪 (1991) 「地方自治体における影響力の計量分析」, 小林良彰編 『政治過程の計量分析』, 芦書房
- Beck, N., Katz, J. N. and Tucker, R. (1998). Taking Time Seriously: Time Series Cross Section Analysis with a Binary Dependent Variable. *American Journal of Political Science*. 42 (4), 1260-1288.
- Berry, F. S. and Berry, W. D. (1990). State Lottery Adoptions as Policy Innovations: An Event History Analysis. *American Political Science Review*, 84 (2), 395-415.
- Berry, F. S. and Berry, W. D. (1992). Tax Innovation in the States: Capitalizing on Political Opportunity. *American Journal of Political Science*, 36 (3), 715-742.
- Blossfeld, H. P. and Rohwer, G. (2001). *Techniques of Event History Modeling: New Approaches to Causal Analysis, 2nd edition*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Dye, T. R. and Robey, J. S. (1980). Politics versus Economics: Development of the Literature on Policy Determination. In T. R. Dye & V. Gray (eds.), *The Determinants of Public Policy*. Lexington, MA: D. C. Heath and Company, 3-17.
- Gray, V. (1973). Innovation in the States: A Diffusion Study. *American Political Science Review*, 67, 1175-1185.
- Gray, V. (1994). Competition, Emulation and Policy Innovation. In C. D. Lawrence & J. Jillson (eds.), *New Perspectives on American Politics*. Washington, D.C.: Congressional Quarterly Press, 230-248.
- Knoke, D. (1982). The Spread of Municipal Reform: Temporal, Spatial and Social Dynamics. *American Journal of Sociology*, 87, 1314-1339.
- Masuyama, M. (2002). Bandwagon Competition or Duration Dependence? Comment on 'Shaping Policy Diffusion'. *Japanese Journal of Political Science*, 3 (2), 265-269.
- Mintrom, M. (1997). Policy Entrepreneurs and the Diffusion of Innovation. *American Journal of Political Science*, 41 (3), 738-770.
- Pollitt, Christopher, and Bouckaert, Geert (2004) *Public Management Reform: A Comparative Analysis*. 2nd ed. Oxford: Oxford University Press.
- Reed, Steven (1986). *Japanese Prefectures and Policymaking*. Pittsburgh, PA: University of Pittsburgh Press.
- Rogers, E.M. (1995). *Diffusion of Innovations, 4th edition*. New York, NY: Free Press.
- Singleton, R., Straits, B. C., Straits, M.M. and McAllister, R.J. (1988). *Approaches to Social Research*. Oxford University Press.
- Steiner, K. (1980). *Progressive Local Administrations: Local Public Policy and Local-National Relations*. In Steiner K., Krauss E. S. & Scott S. C. (eds.), *Political Opposition and Local Politics in Japan*. Princeton, NJ: Princeton University Press, 317-352.

- Tolbert, P. S. and Zucker, L.G. (1983). Institutional Sources of Change in the Formal Structure of Organizations: The Diffusion of Civil Service Reform, 1880-1935. *Administrative Science Quarterly*, 28, 22-39.
- US General Accounting Office (GAO)(1990). *Case Study Evaluations*. GAO/PEMD-10.1.9. Washington, D.C., November.
- Walker, J. L. (1966). The Diffusion of Innovations among the American States. *American Political Science Review*, 63, 880-899.
- Weiss, C. H. (1998). *Evaluation: Methods for Studying Programs and Policies*, 2nd ed. Upper Saddle River, NJ: Prentice Hall.
- Yamaguchi, K. (1991). *Event History Analysis*. Newbury Park, CA: Sage.
- Yin, R.K. (1994). *Case Study Research: Design and Methods*, 2nd ed. Newbury Park, CA: Sage.

(2006.2.10受理)

An Analysis of Diffusion and Developmental Process in Local Governments' Evaluation and Performance Measurement

Shun'ichi Furukawa

Harumi Morikawa

University of Tsukuba

Ibaragi Prefecture

furukawa@sk.tsukuba.ac.jp

morikawa@ipu.ac.jp

Abstract

This paper analyses the diffusion and development process of evaluation and performance measurement in Japan's local governments. We indicate that unprecedented is the analysis of such process focusing on contributing factors to the diffusion and developmental process, and employs event history analysis of 47 prefectures and complementarily a case study approach for 10 representative local authorities. The results of these two approaches suggest three points: first, an eagerness of governors or mayors, a tradition of administrative reform, and change of administration contribute to the initiation and developmental process; second, social-economic factors do not contribute so much as is generally expected, and third, reference group has not impacted on the diffusion process, but a countrywide catching up/imitating strategy among local governments has proved an important contribution to this diffusion. To sum up, political factors are significant, in addition to this strategy.

Keywords

local governments, evaluation, performance measurement, policy diffusion, event history analysis, case study

【研究ノート】

A Study of Evaluation Methodology for Donor-funded School Construction Projects in Developing Countries – through Establishing Evaluation Guidelines –

Satoshi Morita, Yasuyuki Sagara

Graduate School of Agricultural and Life Sciences, the University of Tokyo
aa07181@mail.ecc.u-tokyo.ac.jp

Abstract

This paper reconsiders evaluation methods for donor-funded SCPs at the primary and secondary education levels in developing countries and improvement of future prospects for education assistance. Through literature reviews and interviews with those concerned, this paper examines recent characteristics of SCPs in Japan's ODA, and presents a list of indicators related to SC and conceptual models of assistance framework in education cooperation based on stakeholders analysis. An examination of the actual state of SC in Indonesia as a case study reveals recent trends in developing countries. Finally, evaluation guidelines for donor-funded SCPs in developing countries are established, and key issues are observed, such as: 1) more emphasis on relevance of the means of assistance for SC; 2) the importance of applying the DAC's five criteria from holistic viewpoints, considering stakeholders' roles; and 3) the necessity of considering recently established local standards regarding SC in developing countries.

Keywords

outcome/output indicators, assistance framework, local standards, stakeholders analysis, The DAC's Five Criteria for Evaluation

1 . Introduction

After the World Conference on Education for All in 1990, an international focus on the importance of assistance in basic education has been established, and the number of school construction projects (SCPs) implemented or funded by donor agencies has increased. SCPs are of vital importance for two reasons; 1) the existence of school buildings or classrooms can directly and/or indirectly relate to major concerns in basic education¹; and 2) school

construction (SC) occupies a large part of the total amount spent on education assistance.

Japan is one of the largest donor countries that have provided assistance for SC in numerous developing countries. In Japan's Official Development Assistance (ODA), most SCPs have been conducted in the framework of General Grant Aid (GGA)². Regarding the SCPs in GGA, some donors have made the criticism that the construction costs per school building (or classroom) are much higher than those in other donors' or NGOs' projects³. Positive character-

istics in SCPs in GGA are also pointed out, such as: 1) school buildings are designed to be sustained for 30 to 50 years from the viewpoint that safety and durability are emphasized; and 2) school buildings are constructed even in areas where construction is technically difficult for the purpose of technology transfer to local contractors (IDJ 2004). Thus not only the construction costs but also the viewpoint of whether assistance is provided while meeting the needs of target areas and groups should be considered in evaluation.

Even though SCPs are classified as cooperation in hardware, it is crucial to look at the effectiveness of projects through evaluation from the standpoint of not only whether school buildings are utilized as the basis for education activities, but also whether projects have contributed to the development of local human resources through the school buildings utilization. In that sense, SCPs are unique. Meanwhile, there have not been many studies of evaluation of SCPs that clearly focus on this uniqueness.

2 . Objectives and Methods

The main objectives of this paper are to reconsider evaluation methods for SCPs at the primary and secondary education levels and to improve future prospects for education assistance by establishing evaluation guidelines for donor-funded SCPs in developing countries in an inductive manner. Chapter 3 deals with key issues regarding SC in developing countries by widely examining characteristics of SCPs and their evaluation in Japan's ODA. Chapter 4 focuses on Indonesia as a case study in evaluation of SCPs, using conceptual models based on stakeholder analysis in order to look at the current status and trends in developing countries. Chapter 5 presents evaluation guidelines for donor-funded SCPs in developing countries⁴, and Chapter 6 includes considerations for education assistance.

Table 1 shows the list of projects and types of documents under review. It includes the 14 SCPs in GGA conducted in the fiscal year 2004 and 4 out of a total of 9 in Loan Aid since 1990⁵ in Japan's ODA, and the WB and ADB projects since 1990 in Indonesia which have a SC component. Also, two documents published by the Ministry of National Education of Indonesia (MoNE) are referred to later.

3 . SCPs in Japan's ODA

3.1 General Description of " SC "

Based on review results, " school construction " can be generally described as construction⁶ or renovation of school buildings (including classrooms) at general education institutions at the primary and secondary education levels. SCPs also involve, in many cases, construction of principals' offices, teachers' rooms, storage rooms, libraries, laboratories, multi-purpose rooms, playgrounds, and support facilities including water supply, electrical systems and restrooms. Furthermore, soft components for school management and the development of pilot models, and provision of school furniture and equipment are often involved.

3.2 Expected Effects and Indicators in SCPs

Table 2 shows summarized analysis results of expected effects from SCPs and corresponding key indicators⁷. Direct effects are generally divided into two categories, while indirect effects are divided into three categories. Each category contains several items. The items marked with circles are actually examined, while the rest are not covered or even mentioned in the ex-post evaluation reports. Only a few indicators are found for about half of the items, and no indicator is found for one of them. Not all indicators in the list are directly related to the construction of school and support facilities. Only those underlined are actually employed in the ex-post evaluation reports.

Table 1 List of Documents under Review

Japan's ODA (GGA)			
Countries	Projects	B/D Study Report	Ex-ante Eva. Table
Afghanistan	School Construction	○	-
Cameroon	Construction of Primary Schools (Phase III)	○	○
Cambodia	Construction of Primary Schools in Phnom Penh	○	○
Zambia	Construction of Basic School In Lusaka (Phase II)	-	○
Nigeria	Construction of Primary Schools	○	○
Nicaragua	Rehabilitation of School Facilities of Basic Education in the Department of Managua	-	○
Niger	Construction of Primary Schools in the Regions of Dosso and Tahoua	-	○
Vietnam	Improvement of Facilities of Primary Schools in Northern Mountain Region (Phase II)	○	○
Madagascar	Construction of Primary Schools (Phase II)	○	-
Mali	Construction of Primary Schools (Phase II)	○	○
Mongol	Improvement of Primary Education Facilities (Phase III)	○	○
Laos	Construction of Primary Schools	-	○
Lesotho	Construction of Primary Schools	○	○
East Timor	Reconstruction of Primary and Junior Secondary Schools	○	○
Japan' ODA (Loan Aid)			
Countries	Projects	Ex-ante Eva. Table	Ex-post Eva. Report
Jordan	Education Sector Loan	-	○
Philippines	Primary Education	-	○
Indonesia	Junior Secondary School Buildings Construction	-	○
Morocco	Local Junior Secondary Schools Improvement	○	-
ADB and WB Projects in Indonesia			
Donors	Projects	Appraisal Report	Ex-post Eva. Report
Asian Development Bank (ADB)	Senior Secondary Education	-	○
	Private Junior Secondary Education	-	○
	Basic Education	-	○
The World Bank (WB)	Junior Secondary Education* 1) East Java and East Nusa Tenggara; 2) Central Indonesia; 3) Sumatra	○	-
	Basic Education* 1) West Java; 2) Sulawesi and Eastern Islands; 3) Sumatra	○	-

Source: The authors, based on donors' URLs

Note: * there are three different projects under each of the same listed titles.

4 . SC in Indonesia Case

4.1 Comprehensive Approach in Education Cooperation

SC has been merely one (sub) component of the projects that ADB and the WB have conducted in the education sector in Indonesia. Figure 1 compares three conceptual models of assistance framework in education cooperation comprising a SC component.

All three show how capital, as input, finally manifests effects on students as an end outcome. In the GGA model, in which a Japanese consulting firm plays an important role as a contractor of (A) grant, two main activities are basically depicted: (a) school buildings and facilities construction as cooperation in hardware; and (b) the soft component of school management support through the central/local government and the school committee. Also, (c) the effects of technology

Table 2 Expected Effects from SCPs and Corresponding Key Indicators

Items	Key Indicators
1. Direct Effects	
(1) Improvement of Education and Study Environment	
<input type="radio"/> Expansion of student capacity	1, 2, 3, 4, 5, 18, 20, 25
<input type="radio"/> Improvement of access to education	1, 2, 8, 9, 18, 19, 25
<input type="radio"/> Improvement in the quality of study	6, 7, 20, 21, 22, 25, 26
Improvement of sanitary environment at school	11, 12, 13, 18, 23
<input type="radio"/> Improvement of school buildings/classrooms	10, 16, 25
(2) Improvement of School Management	
<input type="radio"/> Maintenance of school buildings and support facilities	14, 17
Improvement of the staff' s capacity (soft components)	24, 25
Establishment of maintenance systems for education materials and equipment (soft components)	24
<input type="radio"/> Improvement of school management systems and plans (soft components)	16, 20, 24, 25
2. Indirect Effects	
(1) Secondary Effects on Target Society/Promotion of Community Activities	
<input type="radio"/> Facilities use for adult education activities, teachers in neighboring areas and community activities	3, 4, 5, 6
Promotion of community' s participation in school management and maintenance	24
Improvement of sanitary environment around school	11, 12
Expansion of employment opportunities	28
<input type="radio"/> (2) Reduction of Facilities and Equipment Maintenance Fees	15, 27
(3) Effects of Technology Transfer to Local Contractors	(N.A.)**
3. Other (Long Term Effects)	
<input type="radio"/> (1) Correction of Discrepancies in Education Quality among Regions*	2, 3, 21, 22, 26
Key indicators*** directly related to construction of school and support facilities	
1. enrollment rate in target regions	
2. number of students at target schools	
3. number of (increased) classrooms at target schools	
4. average size of classroom at target schools	
5. frequency of classroom use at target schools	
6. frequency of (newly established) school facilities use at target schools	
7. sufficiency of school furniture and education materials at target schools	
8. average commuting time at target schools	
9. average attendance rate and hours at target schools	
10. percentage of temporary or old unsafe classrooms in target regions	
11. percentage of schools in target regions at which safe water can be obtained	
12. percentage of schools that have restrooms in target regions	
13. ratio of students per stall in restroom at target schools	
14. ratio of schools in target regions that have a principal' s office, a teachers' room and a storage room	
15. amount of school building rent at target schools	
16. conversion rate from temporary to permanent school buildings at target schools	
17. five scales rating of facilities maintenance (e.g. durability, water and electricity use, rain leakage, cracks in walls, and school furniture installment)	
Other related indicators***	
18. percentage of female students in target regions	
19. dropout rate in target regions	
20. conversion rate from part-time to full-day school in target regions	
21. ratio of students to teacher at target schools	
22. number of (qualified) teachers in target regions	
23. number of active school committees for sanitary management at target schools	
24. number of activities implemented according to the school committee' s activities plans at target schools	
25. degree of student satisfaction with education services at target schools	
26. average scores of nation-wide exams in target regions	
27. tuition per student at target schools	
28. employment rate in target regions	

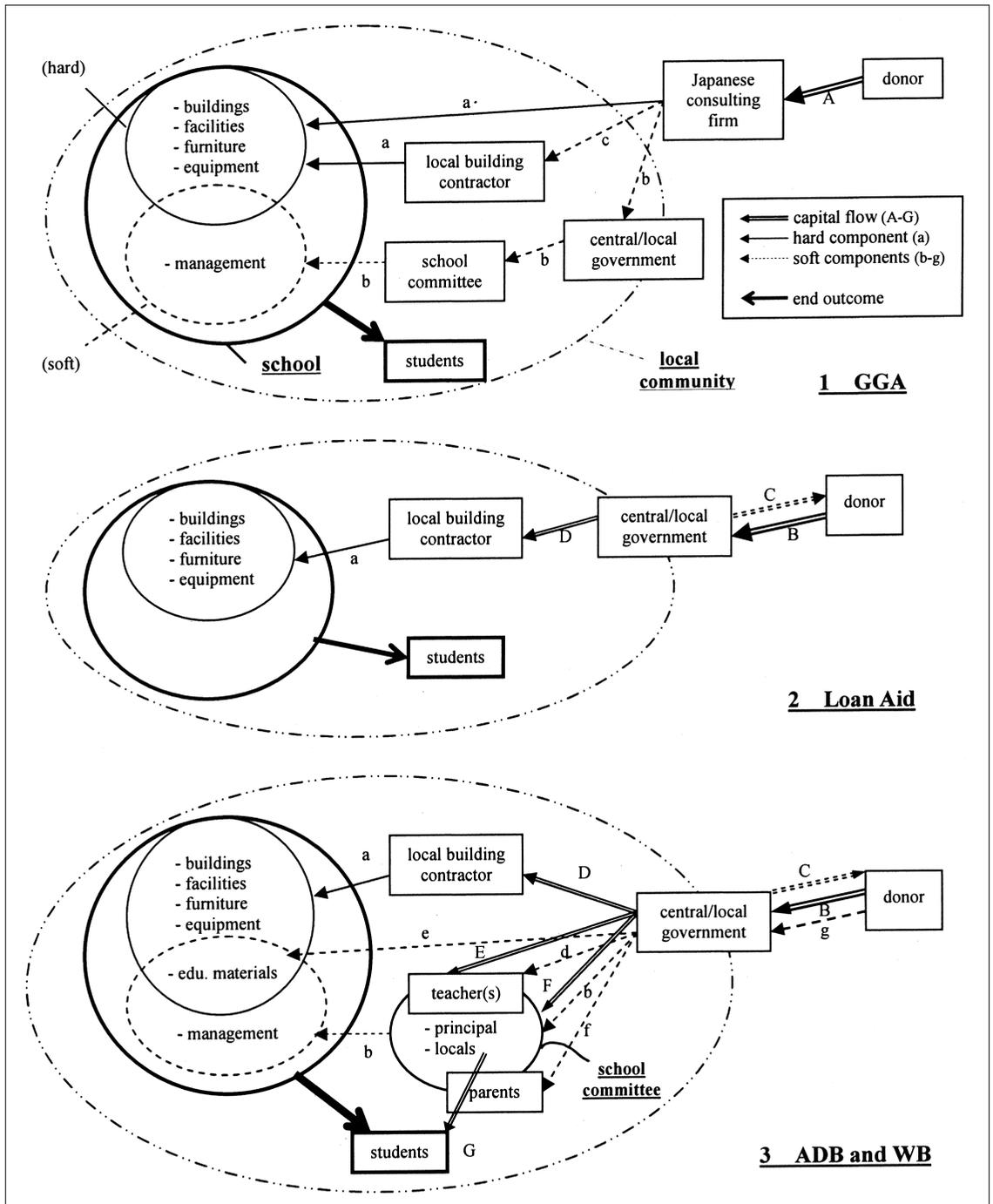
Source: The authors, based on review results of B/D Study reports, ex-ante evaluation tables and ex-post evaluation reports

Note: * in the case of nation-wide projects in which some specific indicators are used for comparison among target regions

** no indicator is found in the reviewed documents.

*** underlined indicators are actually employed in the ex-post evaluation reports.

Figure 1 Three Types of Conceptual Models of Assistance Framework in Education Cooperation



Source: The authors, based on review results

Note: In reality, there may be some exceptional cases of projects to which these models cannot be completely applied.

transfer are depicted. In the second model of Loan Aid in Japan's ODA, SC is the main component under (B) the loan, in which the recipient country uses more discretion and involves fewer stakeholders. Also, (C) repayment of loan is assumed after project completion.

Meanwhile, in the third model of the two agencies' framework, in addition to (a) and (b), other components through the central/local government are shown, such as: (d) training of teachers; (e) development and provision of materials for teaching; (f) provision of information to students' parents; and (g) other possible components. Also, (D) capital for construction; (E) salary for teachers; (F) school management budget; and (G) scholarships for students are important elements. This model is more comprehensive regarding which interaction among more stakeholders within the community can be seen.

Accordingly, the two agencies' evaluation reports do not focus only on the SC component, but rather cover all components widely. One issue is noteworthy: the connections between components are not always clear and thus, attribution of project effects to activities is not precisely described. Therefore there is still room for improvement in their evaluation i.e. focusing, in order to deepen analyses, on different aspects of a project in the application of the DAC's five criteria from holistic viewpoints. This approach can also be applied in SCPs in GGA.

4.2 The Government's Evaluation Framework and System

"New School Units Development Program Implementation Guidebook," issued by MoNE in 2004, was originally developed with the WB's support, for the purpose of smoothly implementing, through provision of block grants, new SCPs at the lower secondary level⁸. From review results, the guidebook illustrated key issues, including: 1) specific criteria for the selection of school sites; 2) a detailed description of roles and responsibilities of stakeholders who are involved in the project implementation and management, including the MoNE and local government officers at each administrative level,

members of the SC committee and local consultants in the field; and 3) a clear range of project monitoring and evaluation, and description of persons in charge and their tasks. In other words, the government owns its framework and analysis viewpoints for evaluation of SCPs at the lower secondary level in the local context. This is a case example that proves that local standards and criteria for management and evaluation of SCPs exist in a developing country.

"School Improvement Grant Program II SC Advisory Services," also issued by MoNE in 2004, mainly focuses on technical aspects in renovation of elementary and junior high schools through the second school improvement grant program by the Indonesian Government, and deals with renovation activities through the program at 2,875 schools in 16 provinces in 2003. From review results, the report illustrated general key issues regarding evaluation of management and results of grant programs, including: 1) quality and efficiency of renovation activities; 2) effects of grant; 3) transparency of activities; and 4) social participation. The report also deals with specific points, such as: 1) aesthetic elements as a construction quality issue; 2) the importance of a higher ratio of capital allocation for hard components as well as a provision of manpower or donation from the local community; and 3) the importance of a higher degree of understanding and recognition of tasks and responsibilities at all administrative levels for promoting smooth implementation of activities and preventing people from concealing budget information. Indonesia obviously has its own evaluation system for school renovation activities with clear recognition of some issues specific to school renovation in the field.

The question may be asked: is Indonesia considered to be a special case among developing countries? In fact, there are other countries that have recently established similar kinds of guidelines with donor agencies' support. In Afghanistan, the Ministry of Education and UNICEF have developed guidelines for school facilities construction and published them in 2003 (JICA *et al.*2004). In East Timor, the basic

designs for education facilities established in the Fundamental School Quality Project, supported and initiated by the WB in 2002, have been adopted as the standards (JICA and MKJ2003). Likewise, in Vietnam, the designs for school buildings used in a Japanese GGA project have been adopted as the standards (IDJ2004). These countries, including Indonesia, are not considered to be the only cases.

5 . Evaluation Guidelines for SCPs

5.1 Issues in Current Evaluation of SCPs

One main issue in current evaluation of SCPs that emerged from review results was the difficulty in using the DAC's five criteria effectively. In some cases, evaluation viewpoints and items are confused and corresponding indicators are insufficiently set or not employed, as mentioned in Chapter 3. In other cases, holistic viewpoints are lacking in evaluation of projects that are composed of multiple components involving different stakeholders, as stated in Chapter 4. Another issue is new trends in the establishment of guidelines and local standards for SC and relevant systems in developing countries, as discussed in Chapter 4. Considering these points, an extract of newly established evaluation guidelines for SCPs is presented in this chapter.

5.2 Premises for Guidelines

The guidelines can be applied for evaluation of SCPs at any time, regardless of whether the target projects, as described in 3.1, are completed, on-going or in the future⁹. Potential users are mainly those involved in evaluation of SCPs in developing countries, which are supported by any donor regardless of the type of assistance framework. The guidelines are developed with the application of the DAC's five criteria, and from the standpoint that extended effects on school management or local communities from other soft components or support facilities construction should also be examined.

Another important premise is to integrate the recipient country's existing system. Evaluators need to

obtain relevant information in the recipient country, and to examine whether existing guidelines and local standards for SC have been applied in project formulation and implementation, and if not, the reasons why. In the case that the target country has its own established evaluation system, the possibility of utilizing existing reports or data and applying local viewpoints in evaluation may be considered.

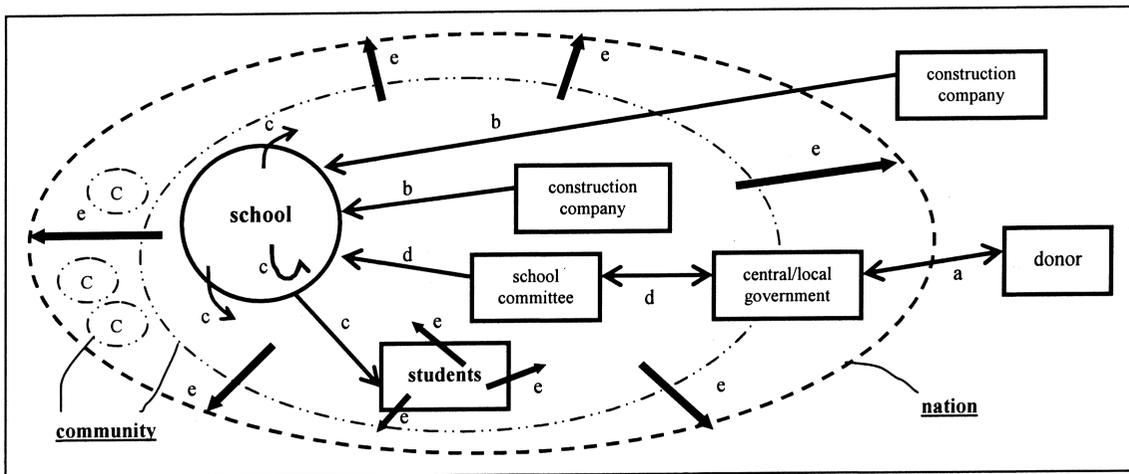
5.3 Guidelines Based on DAC's Five Criteria

Figure 2 shows a conceptual picture of the five criteria and relationships among stakeholders. This has been developed, based on the discussion in 4.1, for the purpose of visualizing main viewpoints for the criteria linked to each relationship¹⁰. The main viewpoints are as follows: (a) relevance principally relates to the policies between the donor and the recipient country; (b) efficiency directly concerns the performance of the construction company itself; (c) effectiveness usually refers to positive changes brought to school, students and the target community from SC and other components; (d) sustainability particularly focuses on school management issues and the school committee's role; and (e) impact basically deals with the issue of human resources development in the community, and sometimes with extended effects at the national level if the project is widely conducted in other regions. Detailed evaluation viewpoints for each criterion are described as follows.

5.3.1 Relevance

Relevance is assessed from the main viewpoints 1 to 3 as shown in Table 3. Key issues for each viewpoint are shown in *italic*. Regarding the appropriateness of the selection of assistance framework in the case of bilateral aid, which has never been discussed in the actual evaluation of SCPs in the past, differences in construction costs per school building within the selected framework may be questioned. This issue needs to be discussed from the recipient country's viewpoint within their policy framework. In fact, the construction cost per building is merely an indicator that could show the durability of

Figure 2 Conceptual Picture of DAC's Five Criteria and Relationships among Stakeholders



Source: The authors

Table 3 "Relevance" and Evaluation Viewpoints

	Relevance
Viewpoints and Key Issues	1. Consistency with the recipient country's national policy in the education sector ○ <u>Coordination with other donors</u>
	2. Needs of target groups and regions in the recipient country ○ <u>Appropriateness of the selection of assistance framework</u> • (In Japan's ODA case) Loan Aid, Grant Aid (GGA, etc.), Technical Cooperation • Assistance approaches (e.g. participatory approach) • Consideration of the influences caused by natural disasters (trade-off between the construction cost per school building and the number of school buildings) ○ <u>Coordination with other assistance frameworks</u> • (In Japan's ODA case) Coordination with Loan Aid, Grant Aid and Technical Cooperation ● <u>Appropriateness of selected target areas and groups</u> ● <u>Consideration of the environment</u> ● <u>Lack of unrelated profitability</u>
	3. Consistency with donor country's aid policy, and its relative development level* in target region ○ <u>Consideration of the construction technology level</u>

Source: The authors, based on relevant reports (same for Table 4 to 7)

Note: (○) newly emerged issues in this study; (●) already observed issues in past evaluation reports (same for Table 4, 6 and 7)

* only in cases where the project is conducted by a bilateral aid agency

the building. The degree of durability should be determined considering the possibility of being affected by natural disasters in target regions. Accordingly, the construction cost per building and the number of schools to be constructed may change. That is, this cost issue highly involves political and technical judgment based on the recipient country's

request. Thus construction costs are more a matter of the relevance of the target project, rather than the efficiency.

In terms of the issue of the donor country's relative level of development in the target region, evaluators need to look at the degree to which the construction technology level in the region has been

Table 4 “ Efficiency ” and Evaluation Viewpoints

	Efficiency
Viewpoints and Key Issues	<p>1. Efficiency within the selected assistance framework</p> <ul style="list-style-type: none"> ○ <u>Appropriateness of original plans in the local context</u> <ul style="list-style-type: none"> • Project range and its quality (decision-making or selection of construction designs – e.g. the size and type of school buildings, aesthetic elements – and materials, construction of support facilities, and provision of school equipment) • Project designs (budgets, project period, implementation system and labor force) ● <u>Differences between original plans and actual results</u> <ul style="list-style-type: none"> • Project budgets • Project period • Purchase and/or selection of equipment and manpower (including the employment situation of consultants, and performance of members of school committee and contractors)

considered in decision-making in terms of coordination with other assistance frameworks (especially technical cooperation) and implementation of soft components within the target project. The issues of coordination with other donors and assistance frameworks should also be considered in discussing relevance of the target project.

5.3.2 Efficiency

The issue of comparison of the costs per school building within different assistance frameworks is not a matter that concerns efficiency. Rather, differences between original plans and actual results within the selected assistance framework should be assessed regarding particular items shown in Table 4, and reasons questioned on the assumption that the selected framework is appropriate. The guiding premise here is that the original plans, including “ project range ” and “ project designs, ” were appropriately made. As an important basis for value judgments, local knowledge and systems should be emphasized, which is an important evaluation viewpoint that has emerged from this study. Regarding “ project range, ” evaluators should examine the appropriateness of the original plans based on the technical knowledge of local technicians or consultants in the region, considering local construction criteria or the current status of equipment purchase. In terms of “ project designs, ” items related to performance of stakeholders are assessed with maximum consideration of existing

relevant guidelines or rules in the country regarding stakeholders' tasks and responsibilities in SC.

5.3.3 Effectiveness

Firstly, construction of school buildings, laboratories, playgrounds and support facilities has effects¹¹ on target groups or societies. These effects are categorized as: 1) direct effects i.e. improvement of education and the study environment for students; and 2) indirect effects on local communities. Effects from soft components are also examined if they have been involved in the project range. Key issues shown in Table 5 are all re-categorized from the contents of Table 2, and the corresponding indicators for each issue in the table can be employed as they appear. To assess direct effects, evaluators first need to obtain existing data or conduct surveys to collect necessary data in the field, and then consider the possibility of applying existing local criteria or numerical values, with a view towards gender equality. Meanwhile, it is generally difficult to precisely measure effects on quality of education. Among the items listed under indirect effects, the issue of technology transfer is a new important consideration in the evaluation of SCPs. This is true, not only in GGA, but in Japan's ODA case, regardless of the type of assistance framework, or even in other donor's assistance frameworks. From the observation in 3-2, it is desirable to develop indicators regarding this matter.

Secondly, it is important to examine effects from

Table 5 “ Effectiveness ” and Evaluation Viewpoints

	Effectiveness
Viewpoints and Key Issues	<p>1. Effects on target societies or groups from construction of school buildings and facilities</p> <p><Direct Effects></p> <ul style="list-style-type: none"> ◎ <u>Improvement of education and study environment for students</u> <ul style="list-style-type: none"> ▪ Enrollment situation for students and school environment ▪ Quality of education ▪ Safety and sanitary environment at school <p><Indirect Effects></p> <ul style="list-style-type: none"> ◎ <u>Promotion of activities in local communities</u> <ul style="list-style-type: none"> ▪ Degree of facilities utilization by local teachers, or for education activities in local communities ◎ <u>Usefulness of support facilities in local communities</u> <ul style="list-style-type: none"> ▪ Level of restroom and water facilities use in local communities ▪ Sanitary environment around school ◎ <u>Effects of technology transfer to local contractors</u> ◎ <u>Expansion of employment opportunities</u> ◎ <u>Reduction of maintenance costs for facilities, school fees and commute fees*</u> <p>2. Effects on target societies or groups from soft components**</p> <ul style="list-style-type: none"> ◎ <u>Improvement of utilization management of school buildings and facilities</u> <ul style="list-style-type: none"> ▪ Maintenance of principal's office, teachers' room, equipment, and storage areas ▪ Local community's understanding, cooperation and involvement

Note: (◎) re-categorized issues from the contents of Table 2 (same for Table 6)

*these can be particular issues for target regions or project contents; ** only in cases where soft components are involved

assisting the management of daily utilization of school buildings and facilities as soft components. In this case, the local community's participation is key i.e. the actual level of the local community's understanding, cooperation and involvement.

5.3.4 Sustainability

Sustainability is assessed from the three main viewpoints as shown in Table 6. Among the key issues listed, the appropriateness of conditions for school location is an important consideration that has emerged from this study. Evaluators need to assess the neighborhood environment while considering the possibility of applying existing guidelines or criteria set by the recipient country. Another new issue in the evaluation of SCPs is the degree to which school buildings and support facilities satisfy geological, meteorological and socio-cultural conditions. This should be examined considering local criteria based on technical knowledge of local technicians and consultants. Listed items should be examined

periodically after the completion of the project.

5.3.5 Impact

Mid-/long term project effects on the education sector in target regions or the recipient country should be examined. More specifically, as shown in Table 7, evaluators need to assess the degree to which the project contributes to the development of human resources and the education environment in the region, or at the national level in cases where the project is conducted nation-wide. Indicators such as the enrollment rate and the number of students per class can be used, and changes in types of employment or future study of graduates can be assessed. Also, correction of discrepancies in education quality among regions over a long term is another consideration. Evaluators, however, should keep in mind that it is very difficult to scientifically prove direct project effects even when those indicators are used.

In cases where a pilot component is involved using model schools in the project, it is necessary first

Table 6 “ Sustainability ” and Evaluation Viewpoints

	Sustainability
Viewpoints and Key Issues	<p>1. Appropriateness of school and facilities maintenance</p> <ul style="list-style-type: none"> ◎ <u><i>Appropriateness of maintenance system and plans</i></u> <ul style="list-style-type: none"> • Status of school and facilities maintenance • Technical standards, capacities and clarity of rights and duties of those in charge and concerned • Maintenance and repair plans of school buildings and support facilities (including budgets) <p>2. Durability and safety of school buildings and support facilities</p> <ul style="list-style-type: none"> ○ <u><i>Appropriateness of conditions for school location</i></u> <ul style="list-style-type: none"> • Harmony with the environment • Safety for students ○ <u><i>Degree to which natural and socio-cultural conditions are satisfied</i></u> <ul style="list-style-type: none"> • Appropriateness in technical aspects <p>3. Appropriateness of school management</p> <ul style="list-style-type: none"> ◎ <u><i>Appropriateness of and future prospects for school management</i></u> <ul style="list-style-type: none"> • School management capacity and clarity of rights and duties of local government officers, school teachers and staff, and members of school committee • Community’s school management capacity or involvement in school management • Key persons’ cooperation and involvement in target regions ● <u><i>Soundness and transparency of financial situation</i></u>

Table 7 “ Impact ” and Evaluation Viewpoints

	Impact
Viewpoints and Key Issues	<p>1. Degree of contribution to the education sector</p> <ul style="list-style-type: none"> ● <u><i>Degree of contribution to the development of human resources in target regions</i></u> <ul style="list-style-type: none"> • Education environment in target regions • Mid-/long term effects on target groups ● <u><i>Degree of contribution to the education sector at the national level*</i></u> <ul style="list-style-type: none"> • Education environment at the national level • Correction of discrepancies in education quality among regions <p>2. Outcome produced from a pilot model component** using model schools</p> <ul style="list-style-type: none"> ○ <u><i>Usefulness of the developed model and its applicability</i></u> <ul style="list-style-type: none"> • Manual of experiences at model schools ○ <u><i>Appropriateness of component designs for applying the model</i></u>

Note: * only in the case that the project is conducted nation-wide; ** only in cases where a pilot component is involved

to distinguish target beneficiaries in the pilot component from other beneficiaries in the project¹², and examine the usefulness of the developed school management model, its potential applicability to other regions, and the appropriateness of component designs for applying the model. These issues have rarely been discussed in actual evaluation reports in the past.

5.4 Remarks for Application of Guidelines

Evaluators should first consider exogenous conditions and their influence on the target project. Main exogenous conditions related to SCPs include: 1) consistent allocation of budgets for school management and facilities maintenance; 2) improvement of infrastructure around schools and access conditions; and 3) employment and placement of teachers. These conditions are directly subject to the

decision-making of the central and local governments in the recipient country. Also, 4) improvement of education regarding sanitation is an important issue related to construction of support facilities.

The contents of statistical data in the education sector of the recipient country highly affect the way of evaluation. Evaluators need to clearly understand the education policy of the target region or country and make evaluation designs utilizing available information or data, and may have to change the designs at any point, depending on the situation.

These guidelines emphasize the importance of local systems and standards in evaluation and will, no doubt, be refined through utilization in the field. In the case that users are identified, evaluation viewpoints can be more specific. It is desirable to add new viewpoints, items and indicators after examining more cases of SCPs. However, some criteria regarding SC that are specific to a certain country, may not necessarily be applied in other countries due to the different socio-economic and cultural factors, and thus, collection of such data may sometimes not be useful. Meanwhile, it is meaningful to collect indicators and relevant data regarding the improvement of education and study environment for students in relation to follow-ups for EFA and the Millennium Development Goals.

6 . Summary and Prospects

This study presented a list of collected indicators related to SC, conceptual models of assistance framework in education cooperation, new trends regarding SC in developing countries, a conceptual picture of the five criteria and relationships among stakeholders for evaluation of SCPs and an extract of newly established evaluation guidelines for SCPs. Key issues that emerged in establishing the guidelines include: 1) the issue of cost per school building should be discussed from the viewpoint of relevance; 2) in order to examine extended effects of SC and other components, it is necessary to apply the five criteria

from holistic viewpoints, considering stakeholders' roles in the local community; and 3) various systems regarding SC are being established in the education sector in developing countries, and evaluators should take them into maximum consideration in evaluating SCPs.

Recently, a shift from single projects in education to comprehensive education assistance has been observed at some donor agencies. Meanwhile, evaluation methods for education assistance may accordingly change to become more systematic (Muta2003). For instance, there has been a move among donor agencies towards joint evaluation in basic education (Imoto2003). The shift of interest in assistance evaluation from the project level to the program and policy levels is another issue that has arisen in evaluation of education assistance (Nagao2003). This shift is key to promoting coordination not only in evaluation, but also in aid implementation among donors. Regarding SCPs in developing countries, donors need to adjust and modify their project designs in accordance with existing local systems related to SC. The number of single projects may decrease and comprehensive education assistance may expand. In that case, evaluation should be implemented in a holistic and systematic way, and the application of the presented guidelines will be helpful.

Acknowledgement

From April 2004 to March 2005, the author (Morita) participated in the research project on " Evaluation of International Education Cooperation Projects, " led by Prof. Masafumi Nagao, at the CICE, Hiroshima University. This paper is based on results of the author's research on establishment of evaluation guidelines for donor-funded SCPs in developing countries, including field surveys in Indonesia. Prof. Koichi Miyoshi at Ritsumeikan Asia Pacific University and anonymous readers, who provided some useful comments, have my deepest gratitude.

Notes

- 1 As seen in the “Dakar Framework for Action” agreed at the World Education Forum in 2000, main issues include: 1) qualitative improvement and quantitative expansion of school attendance in primary and secondary education; 2) reduction of disparities in education; and 3) improvement of educational management (JICA 2002a).
- 2 Since 1979, 184 projects have been conducted in 47 countries, and the total amount provided has been 151.6 billion yen (based on obtained JICA documents).
- 3 The main reasons for this are as follows: 1) contractors are limited to Japanese firms; and 2) the rule of annual budgeting is always applied (IDJ2004).
- 4 Based on results of interviews conducted by the author (Morita), donor agencies in Japan and the WB and ADB have not developed evaluation guidelines for their SCPs. In “JICA Evaluation Handbook for Basic Education Development Projects” published in March 2004, no specific issues or indicators related to SC are mentioned.
- 5 Regarding the other 5 projects, no evaluation documents were published as of December 2004.
- 6 “Construction” of school buildings includes: 1) new construction; and 2) re-construction. The former includes: 1) new construction at a site where a teacher previously regularly offered classes to a certain number of students in the open air; and 2) new construction at a site where no classes of any kind were previously offered. The latter means construction of new school buildings after tearing down originally existing school buildings.
- 7 “Expected effects” emerged mostly from the review results of the B/D study reports, while key indicators are from those of both the ex-ante evaluation tables and the ex-post evaluation reports.
- 8 A block grant is a subsidy that the government has provided since 1998 for constructing new school buildings in areas in poverty or management and maintenance of existing schools, financially supported by the WB or ADB (JICA2001). Prior to this, there was a specially funded program to support construction of elementary school buildings all over the nation, called “Inpres,” initiated in 1973. After achieving a nearly 100

percent average enrollment rate for elementary education at the national level in 1993, the government has raised the goal of realizing a nine-year compulsory education system by 2010.

- 9 Emphasis may be placed on different viewpoints depending on the evaluation timing usually, relevance is emphasized in ex-ante evaluation; relevance and efficiency in mid-term evaluation; relevance, efficiency and effectiveness in terminal evaluation; and sustainability and impact in ex-post evaluation (JICA2002b).
- 10 In evaluation, it becomes important first to place the target project in the recipient country's policy structure. Regarding the relationships between policy structure and the DAC's five criteria, refer to Miyoshi *et al.* (2003).
- 11 The application of economic analysis methods may not be useful in an attempt to measure some of the target project's effects in the education sector. For instance, the cost-benefits analysis is usually not implemented in JBIC projects in the education sector in which there is no profitability in general or benefits cannot easily be measured quantitatively (JBIC2003). The cost-effects analysis involves some technical difficulties regarding data collection, and is also usually not applied (OECF1999).
- 12 Refer to Morita *et al.* (2003) for a model of “pilot project regarding model development and its application.”

References

- Imoto, Sachiko (2003). The Role and Possibility of Joint Evaluation by Aid Agencies - The Case of Joint Evaluation of External Support to Basic Education - . *The Japanese Journal of Evaluation Studies*, 3(1), 3-16.
- International Development Journal (2004). *International Development Journal* (February 2004).
- Japan Bank for International Cooperation (2003). *Training Textbook for Evaluation on ODA Loan Projects*.
- Japan International Cooperation Agency (2002a). *Approaches for Systematic Planning of Development Projects: Basic Education*.
- (2001). *Basic Data on Sector Issues*, Indonesia 2001.

- Vol. 1.
 ---- (2002b). *JICA Evaluation Guidelines*.
 Japan International Cooperation Agency, and Mohri
 Kenchiku Jimusho (2003). *Report on Basic Design Study
 for Project for Reconstruction of Primary and Junior
 Secondary Schools in East Timor*.
 Japan International Cooperation Agency, System Science
 Consultants, Inc., and Fukunaga Architects-Engineers,
 Inc. (2004). *Report on Basic Design Study for Project for
 School Construction in Afghanistan*.
 Miyoshi, Koichi, Y. Hirata, T. Wada, H. Nakazawa, and E.
 Kita (2003). Program Evaluation: The Case of Japan's
 Infections Disease Control Assistance in the Philippines.
The Japanese Journal of Evaluation Studies, 3(1), 43-56.
 Morita, Satoshi, K. Miyoshi, and K. Takamatsu (2003). A
 Study on participatory Secondary Evaluation for the
 Public: The Case of JICA's Evaluation on "Poverty
 Reduction and Rural Development Project in Sulawesi,
 Indonesia." *The Japanese Journal of Evaluation Studies*,
 3(2), 126-143.
 Muta, Hiromitsu (2003). A Case of Comprehensive
 International Cooperation Based on Structural Evaluation.
The Japanese Journal of Evaluation Studies, 3(1), 65-76.
 Nagao, Masafumi (2003). Present Concerns and Issues in
 Evaluation of Educational Cooperation. *Journal of
 International Cooperation*, 6(1), 1-18. Center for the Study
 of International Cooperation in Education, Hiroshima
 University.
 Overseas Economic Cooperation Fund (1999). *Ex-Post
 Evaluation Report on ODA Loan Projects 1999*.
 (2006.1.30受理)

【実践・調査報告】

学生評価インターン制度の拡充に向けた一考察 「ラオスのこども」の組織運営に関する評価調査を事例として

茂木 勇

立教大学大学院

isamsan136@ybb.ne.jp

要 約

我が国の様々な分野で評価が導入・検討されるに伴い、「評価慣行の一層の普及・定着にとって専門性を備えた評価人材の育成が不可欠の課題」(長尾 2005)といった「評価活動を実践・研究する人材」の確保を模索する動きが出てきている。評価人材を中・長期的に確保・育成しようとするとき、これまでの(1)評価実施機関を中心に進められてきた評価人材育成に加え、(2)大学等の高等教育機関がその役割の一部を担うという方向性も検討の余地がある。然るに、本学会研修委員会が展開する評価インターン事業は、後者(2)との関連で、評価初学者に評価実務経験の場を提供する実践的な評価人材育成事業であると言えるが、評価体験が自己目的化する恐れもあり、評価依頼者にとって有意な評価結果をもたらせたかという意味における評価の質は、何れの評価インターンも問われなければならないと考える。そこで本稿では、評価インターンとして国内NGOの組織評価調査に参加した経験をもとに、その評価事例を紹介し、併せてその有用性についてPatton(1997)および長尾(2003)が展開する実用重視評価の理論、とりわけ長尾が掲げる実用重視評価の6つの視点に照らして考察を加える。その結果、いくつかの点において実用重視評価の要素を認めることができたが、同時に課題も見出された。今後の制度拡充のために、評価インターンは実用重視の視点に立ち、結果の有用性に配慮した評価調査を実施することが必要である。

キーワード

実用重視評価、評価結果の活用、組織評価、評価人材育成、学生評価インターン

1. はじめに

我が国における評価研究者及び評価実務者の情報交換の場として、また評価活動の向上と普及のための事業を推進すること(学会規約第3条)を目的に本学会が設立されたのが2000年9月、その後2002年3月に研修委員会が発足し、そして学会設立から3年後の2003年には研修委員会事業とし

て「学生評価インターン出前サービス事業」(以下、評価インターン事業という。)が始動している。

同委員会は当該事業を「学生会員にとっては評価を体験する機会として、委託者である自治体やNGOにとっては外部の評価者による評価を低コストで受けられる機会」と位置づけており、主に評価人材の育成と評価文化の浸透を企図している

(日本評価学会 2005)。

翻って、我が国の評価人材育成を概観するとき、これまでもODA評価に関連するものとして、JICAのPCM研修をはじめ、国際開発高等教育機構 (FASID) や国際開発センター (IDCJ) 等が実施する研修プログラムの例があり、また政策評価ないし行政評価に関しては、総務省の政策評価統一研修や政策評価フォーラムのほか、地方自治体においても評価実務研修が実施されるなど、その役割は主に評価実施機関が担ってきた経緯がある。

一方で、我が国において評価活動が浸透し、様々な分野で評価が研究・実践されるに伴い、「評価慣行の一層の普及・定着にとって専門性を備えた評価人材の育成が不可欠の課題」(長尾 2005) といった指摘も出ている。

それゆえ今後、評価人材を中・長期的に確保していこうとするならば、上述の(1) 評価実施機関を中心としたこれまでの評価人材育成に加えて、(2) 大学等の高等教育機関がその一翼を担うという方向性も検討の余地がある。米国を始めとする諸外国の事例¹を見ると、また我が国においても評価論講座を開講する大学が近年増えつつあることを鑑みるに、その実現可能性は少なからずあると思われる。

然るに、評価インターン事業は、研修委員会の企図によれば、後者(2)の視点との関連で、評価初学者に評価実務経験の場を提供する、優れて実践的な評価人材育成事業であると言えるだろう。ただし、ここで注意が必要なのは、評価依頼者すなわちインターン受入先にとって意味ある介入とならなければ、評価実務としては価値がない点である。したがって、単なる“評価体験”に終始することなく「評価依頼者にとって如何に有意な評価結果をもたらせたか」という意味における「評価の質」は、何れの評価インターン参加者についても問われなければならないと思われる。

以上を踏まえ、本稿では、平成16年度評価インターンとして国内NGOの組織評価に参加した筆者の経験をもとに、まず評価インターンが実施した評価調査の一例を紹介する。そのうえで当該調査についてPatton (1997) が提起し、国内では長尾 (2003) が要約・紹介している実用重視評価の

理論に照らし、長尾の掲げる6つの視点を手掛かりに考察を加える。そして考察の結果を踏まえ、今後の評価インターン制度の拡充のための提言ないし課題の指摘を行いたい。

2. 評価インターン調査「ラオスのこども」の組織運営に関する評価調査の概要²

本節では、評価インターン調査の一連の流れ(計画・実施・報告)について、時系列に沿って振り返ることとしたい。

(1) 評価調査実施の経緯と目的

「ラオスのこども」とは

評価対象団体である「特定非営利活動法人ラオスのこども」³ は、1982年にチャントソン・インタヴォン氏(現共同代表)が地域パザーで売れ残った日本語の絵本をラオスのこどもたちに送り始めた活動(「ラオスの子どもに絵本を送る会」)を前身としており、現在では3本の柱となる活動、すなわち、

創作絵本・紙芝居などの出版事業

読書推進運動(図書箱・図書袋配布プロジェクト、学校図書室の整備・運営、教員へのセミナー開催支援)

子ども文化センター事業(8か所の運営、スタッフ研修支援)

を中心に活動を行っている。設立以来、20年の活動期間を経て、2003年5月にNPO法人の認証を取得、法人化に至っている。また、法人化にむけた準備の過程において、任意団体時代に漠然と共有してきた活動の目的・理念を明確にする一方、新たに理事会を設け、意思決定の手順を定めるなど、組織機構の整備を進めてきた。法人化の狙いは偏に「組織の責任を明確にし、信頼性の高い組織運営を確保しつつ、ボランティアによる多様で自発的な活動を実現すること」にあったが、同団体はその部分に不安を抱えていた。

評価調査の流れ

評価調査の主眼は、したがって「ラオスのこど

も」の組織評価を行うことにあるが、評価調査の基本的な流れとしては、概ね「JICA事業評価」で利用されている手法を参考にしている。JICA事業評価では、評価調査を大きく「調査の計画」「調査の実施」「調査の報告」の3つのステップに分けており、また各ステップの作業内容を明確にしていることから（図1）、プロジェクト評価・組織評価の別を問わず、はじめて評価実務に取り組む評価インターンにとって、評価手順を理解するうえで参考になる（国際協力機構2004）。

オリエンテーションと評価仕様書

調査実施に先駆けて平成16年9月3日、学会（研修委員会）、評価委託団体（ラオスのこども）、評価インターン、評価アドバイザー、助成団体担当者（笹川平和財団）、インターン事務局の出席のもと「評価インターン・オリエンテーション」が開催された。

当日の流れとしては、参加者の自己紹介にはじまり、研修委員会より評価インターン事業の趣旨や概要について説明がなされた後、評価アドバイザーの司会により、「ラオスのこども」と評価インターンの間で評価の目的や手順、スケジュール等を明確にするため、評価仕様書（案）を作成した。

しかしながら、オリエンテーションの時間内で

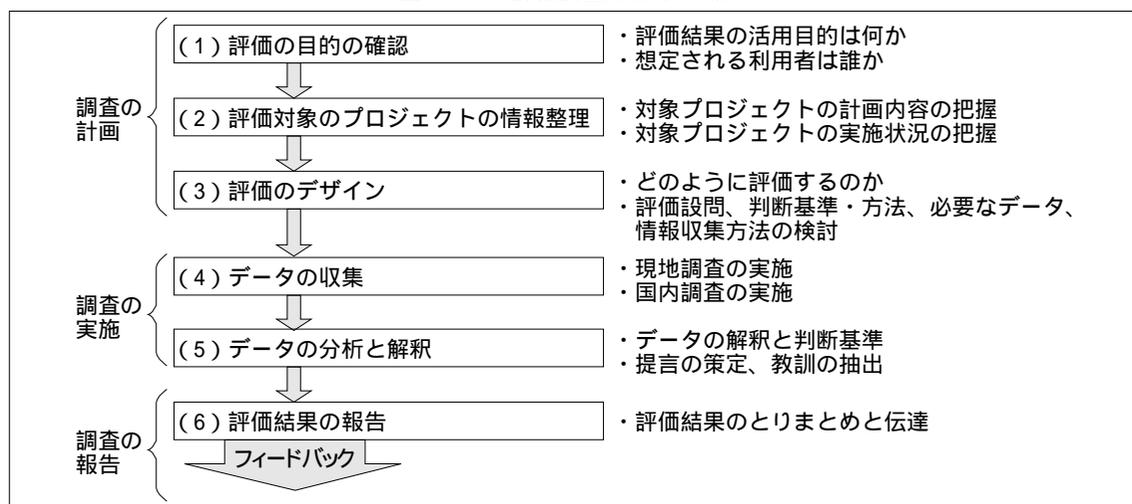
は十分に詰められない部分もあり、この評価仕様書（案）をベースに、「ラオスのこども」と評価インターンの間で数度にわたる意見交換の機会を設けることとなった。その際、評価委託者との「依頼内容の行き違い」を避けるために、評価インターンは以下の3点に留意しながら、出来る限り「対話」の機会を設けて合意形成を図っている。

- （イ）両者のもつ「評価」に対するイメージをより近づけること
- （ロ）団体側の「評価」に寄せる期待を十分に汲みとること
- （ハ）技術的かつ時間的に、評価インターンが担える内容を明確にすること

また、評価仕様書の検討過程で効果的であったのが、「ラオスのこども」の運営メンバー（理事と事務局）を対象としたフォーカス・グループ・インタビュー（以下FGIという。）であった。なぜなら、FGIでは組織の全体像の把握に有効であるのみならず、運営メンバー間において、評価へのイメージや期待に大きな差があったことを評価委託者と評価インターンの双方が認識できたからである。

結果的に、評価仕様書の作成に時間を費やすこととなったが、団体代表者のみならず、運営メンバーから評価に対する印象や期待、日常的な問題意識を共有できたことで、これらを実評価仕様書へ

図1 JICA評価調査の基本的流れ



反映させることができた。このことは、評価結果を利用すると思われる関係者の意向を汲む効果があったと言える。

以上より、今回の評価調査は「ラオスのこども」の組織評価を行うものであり、その目的は、「法人化より1年が経過した同団体について、組織運営に関する全体的な評価を行うこと、加えて積極的な参加を促す、責任体制の明確な組織運営のあり方について改善の選択肢を提示する」ことが確認された。

(2) 組織評価の方法

上記のとおり、「組織運営に関する全体的な評価」を評価目的としたことから、組織運営能力について多角的に分析することとなった。然るに、評価インターンは独自に評価指標を持っておらず、また指標作成も困難であったため、既存の評価指標を使用することとした。文献調査及びインターネットを利用した調査の結果、評価指標候補として以下の4つを掲げた。

特定非営利活動法人ハンガー・フリーワールド
組織評価表⁴

東京ボランティア・市民活動センター「組織評価」資料⁵

Philippine Council for NGO Certification評価指標⁶
国際協力NGOセンター「NGOアカウンタビリティ基準(案)」⁷

これら4つの基準は異なる特徴があったことから、以下の通り比較検討を行っている。

まず4つの基準は大きく と の2つに大別することができた。前者は組織評価のみを扱うもの、後者は組織評価に加えて一部プログラム評価を含むものであり、本調査では前者のタイプが適すると判断した。続いて、 の3つの基準は

の順により詳細な基準となっており、意思決定や情報共有、財務管理といった組織運営を測定する視点(評価項目)をそれぞれ設けており、基本的な要素は何れも共通していると判断した。

これを踏まえて、評価インターンは以下の3つの理由から 国際協力NGOセンター(JANIC)の「NGOアカウンタビリティ基準(案)」(以下、

JANIC基準という。)を指標として採用するに至っている。

第1に、JANIC基準はNGO活動の発展を目的に、NGO関係者が共同かつ自主的に作成した指標であり、評価インターンが使用する評価指標として「ラオスのこども」関係者の納得性が得られやすいと考えたこと。

第2に、JANIC基準を実際に利用した団体に対して、作業における注意事項や実施状況等について聞き取りしたところ「使いやすい基準」との指摘⁸もあり、評価作業が円滑にできるものと考えたこと。

第3に、「ラオスのこども」は、JANIC基準の一部を使って平成16年6月に内部評価合宿を実施しており、すでに同基準の存在を知っていたこと。加えて今回の評価調査は外部評価であるため、団体として過去に実施した内部評価と今回委託した外部評価の比較が可能となり、これも同団体にとっては有益であると考えたこと。

(3) 調査項目とデータ収集方法

JANIC基準はNGOの組織運営に必要とされる要素を9つのカテゴリに分類のうえ、評価項目としている。

【JANIC基準の9つの評価項目】

目的(理念、使命)
意思決定
情報の共有と公開
組織の機構・役割
業務の遂行
財務
明文化
法令等の遵守
組織の強化

「ラオスのこども」はNGOとしての活動歴は長く古参の部類に属するものの、法人化してから1年と間もなく、体制整備を進めている段階であったため、すべての評価項目を評価対象とするか否か、インターンの間でも意見が分かれたが、評価仕様書に立ち返れば「組織運営に関する全体的な

評価」を行う旨を盛り込んでおり、結果として9分野すべてを評価対象項目とした。

次に9つの評価項目ごとに設けられている質問項目（67項目）について、その判断根拠となるデータ内容、対象となる関係者、調査の手段等を一覧化し、明確に判断が行なえるよう様式化した（表1-1、1-2）。

また、評価調査におけるデータ収集については、以下 ~ の方法により行っている⁹。

公開資料・文献レビュー

イベント会場の視察

運営会議の傍聴

内部書類閲覧

聞き取り調査（理事、事務局員、ボランティア、元事務局員）

アンケート調査（メーリングリスト利用）

今回、 ~ の各手法を用意しているが、これは複数のデータ収集方法を組み合わせることでデータの精度をあげようとする「トライアングレーション（三角検証）」¹⁰ の考え方を実践した結果である。また、データ収集方法として掲げたアンケート調査については評価委託者の関心が特に高く、アンケートの質問内容について3度の協議・修正を経て実施するに至った。回収率は低かったものの、評価委託者の意向を十分に汲んだアンケート調査となっている。

（4）レイティング作業

評価作業は、まず9つある評価項目を3つに分け、評価インターン3名がそれぞれ3項目ずつ主担当となった。続いて、各自でひと通りの評価レイティングを行った後、評価者3名で67の質問項目ごとにレイティング値とその判断根拠を示しつつ、評価の「すりあわせ作業」を行った。この「すりあわせ作業」では、主担当を司会に評価インターンがそれぞれのレイティング値を示しながら、根拠となる書類の記載事項、各種規程の存在、聞き取り調査の結果、アンケート集計結果などを材料に協議を重ねる形式をとった。とりわけ評価インターン間で数値に隔たりがある場合は、その判断根拠を相互に示して協議し、再度レイティング

を行った。この方式は多くの時間がかかる一方、インターン3人で協議するため、判断材料の見落としや勘違いを防げるため、評価レイティングの精度を維持する効果があった。

これらの作業を経て、項目別評価結果シートにまとめるとともに、本調査のレイティング値と評価委託者が平成16年6月に実施した内部評価結果を比較するために、チャート比較図を作成して、これを対比させた（図2）。

図2によれば、評価委託者と評価インターンのレイティング値は「2.意思決定」でほぼ一致したものの、「3.情報共有と公開」、「4.組織の機構・役割」、「7.明文化」の3項目では評価インターンの数値の方が高く、またその他5項目では評価委託者の数値が高くなっており、相対的に評価インターンのレイティング値の方が低い傾向にある。ただし、数値の大小よりは、チャート図の形状（凹凸）が異なる点が両者の認識の差であると考えられるから、外部者から「ラオスのこども」を見たとき、情報共有と公開、組織の機構・役割、明文化の3項目に組織運営の強みを見出すことができると判断した。

（5）評価結果に基づく改善策の提案

評価結果に基づく提言・改善策の提示については、内容の目安として「大提言」「中提言」「小提言」の3種類に区分した。

大提言：団体の目的や機構、人員など組織運営の根幹を抜本的に改善するもの
中提言：団体の組織運営において、その機能向上に資する改善を行うもの
小提言：団体の内部事務を改善するもの

本件調査の結果を踏まえ、評価インターンから、大提言2件・中提言4件・小提言6件からなる合計12件の提言・改善策を提示した（表2-1,2-2）。これらの提言・改善策を含む評価結果概要について、平成17年2月に理事および事務局に対して報告し、その後同年8月「ラオスのこども」定期総会¹¹の場においても評価結果概要を報告した。これは「より多くの関係者が集まる定期総会の場で評価結果報告会を実施することで、組織の問題点

表1-1 項目別評価シート（「7. 明文化」のサンプル）

大項目	7. 明文化	項目番号	7-1	担当者	茂木
評価項目	組織の目的（理念、使命）を明文化する。				
達成度	4	指標	団体のミッションは誰からも分かるように明文化されているか		
参考データ	ラオスのこども定款				
調査対象	<input checked="" type="checkbox"/> 理事会 <input checked="" type="checkbox"/> 事務局 <input checked="" type="checkbox"/> 監事 <input type="checkbox"/> 活動会員 <input type="checkbox"/> その他	5名 3名 2名 ()			
調査方法	<input checked="" type="checkbox"/> 書類閲覧 <input type="checkbox"/> ヒアリング <input type="checkbox"/> アンケート <input type="checkbox"/> 現地見学 <input type="checkbox"/> その他				
評価メモ	<ul style="list-style-type: none"> ・定款第3条（目的）において団体の活動目的・使命は明確にされている。 ・また、ラオスのこども通信の裏表紙には定款3条（目的）が記載され、明文化に加えて、常に社会に発信する姿勢が見受けられる。 以上より、評価結果を「達成している（=4）」とした。				

（注）評価項目「7. 明文化」の一つ「7-1. 組織の目的（理念、使命）を明文化する」の評価シートを抽出掲載したもの。

（出所）茂木・山田・山村（2004）

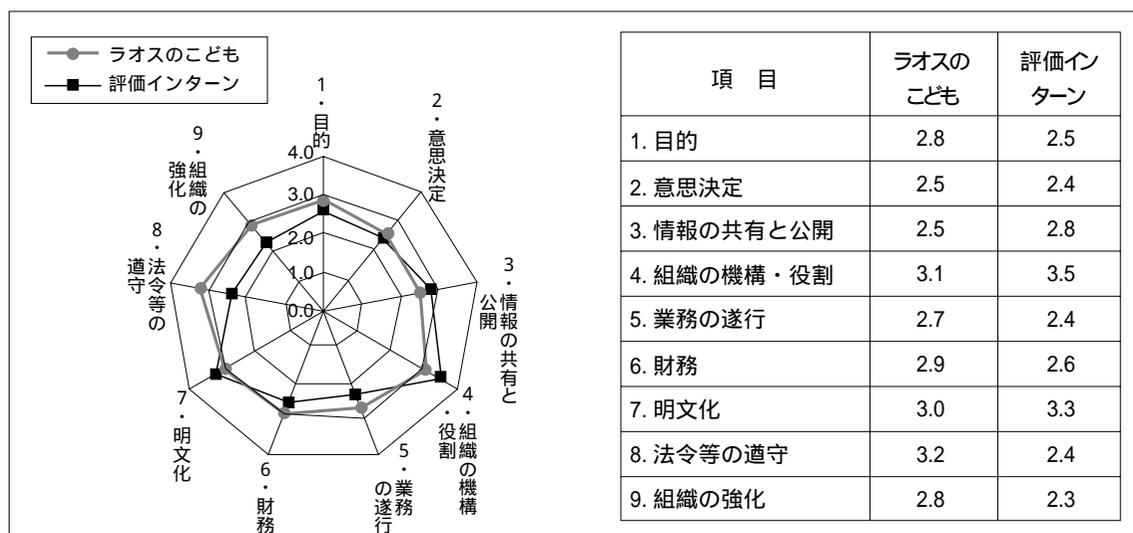
表1-2 「ラオスのこども」組織評価 評価シート（大項目1・2：サンプル）

評価大項目	小項目	到達	判断基準・指標・根拠	必要なデータ	調査対象	調査方法
1. 目的	1-1 組織の目的(理念、使命)を明確にし、必要に応じて見直す。		アンケート	関係者聞き取り定款	理事事務局会員(ML参加者)	聞き取り調査 アンケート
	1-2 関係者の幅広い参加を得て目的(理念・使命)の明文化と見直しを行う。		アンケート・聞き取り調査	NPO立上げ検討資料		
2. 意思決定	2-1 組織の意思を決定するにあたっては、組織内外の関係者の意思を把握し、反映させる。		議事録、聞き取り調査	定款 議事録 ニュースレター	理事事務局活動会員	聞き取り調査 アンケート
	2-2 意思決定にはより客観的な判断ができるように多分野の人の参加を得る。		議事録、聞き取り調査			
	2-3 意思決定の仕組みとプロセス、基準を明確にする。		定款記載、聞き取り調査			
	2-4 意思決定は、透明かつ民主的に行ない、決定の理由を明らかにする。		議事録、聞き取り調査、ニュースレター			
	2-5 決定する事項について利害関係を有する者は、その決定に加わらない。		聞き取り調査			

（注）9つの評価項目のうち「1.目的」「2.意思決定」を抽出掲載したもの。

（出所）茂木・山田・山村（2004）をもとに筆者作成

図2 レイティング集計結果及びチャート比較図



(出所) 茂木・山田・山村 (2004)

表2-1 評価結果に基づく提言・改善策一覧

区分	提言内容	関連評価項目	提言の区分
大提言 (2件)	団体の目的理解・共有化について改善するもの	1 目的	団体の目的や機構、人員など組織運営の根幹を抜本的に改善するもの
	理事・事務局の兼任体制の見直しを行うもの	4 組織の機構・役割	
中提言 (4件)	団体の情報システムを改善するもの	3 情報の共有と公開	団体の組織運営において、その機能向上に資する改善を行うもの
	新人教育担当理事の創設、またはボランティアコーディネーターの設置	4 組織の機構・役割	
	団体の組織強化(人材育成制度)を図るもの	9 組織の強化	
	団体の組織強化(広報部設置)を図るもの	9 組織の強化	
小提言 (6件)	団体の情報共有(ITの高度利用)を改善するもの	3 情報の共有と公開	団体の内部事務を改善するもの
	新入会員用ガイダンスプログラムの開発	5 業務の遂行	
	団体の内部事務(通信事務)を改善するもの	6 財務	
	団体の内部事務(資金管理)を改善するもの	6 財務	
	団体の内部事務(就業規制)を改善するもの	7 明文化	
	団体の組織強化(イベント担当制)を図るもの	9 組織の強化	

(出所) 茂木・山田・山村 (2004) をもとに筆者作成

表2-2 評価結果に基づく提言・改善策（中提言：サンプル）

提言区分	中提言				
提言内容	新人教育担当理事の創設 / ボランティアコーディネーター設置				
問題点	事務局の日常業務の簡素化・効率化を考えると、現状のボランティア（活動会員）のマネジメントに要している負担を軽減すべきと考える。				
改善案	取りうる選択肢としては2つあるものとする。 （1）理事の中から「新人教育担当理事」を指名する。 （2）または理事に限らず「ボランティアコーディネーター役」の人材を育成して、ボランティア（活動会員）のまとめ役、教育担当の任にあたってもらう。				
（参考）関連評価項目					
大項目	4．組織の機構・役割	項目番号	4 1 . B	担当者	山田
評価項目	組織の全体構成と各機関や部署の役割、権限を明確にしておく。				
達成度	3	指標	組織全体の構成と各機関や部署の役割・権限は明確にされているか		
参考データ	定款 職務権限規程				

（出所）茂木・山田・山村（2004）をもとに筆者作成

表3 受入団体による評価

評価項目	具体的な評価の視点	「ラオスのこども」のコメント
評価実施プロセスの妥当性 = 評価の計画、実施の過程は適切であったと思うか。それはなぜか。	（ア）評価対象への理解度	・ 団体内外のいろいろなポジションの人間に聞き取りしていたので組織の全体像について必要なところまで理解したと思われる。
	（イ）調査項目・評価指標の適切性	・ 団体運営を中心的に行うものと、ボランティアや支援者の間でのような意識の差があるのかについて特に関心を寄せていたが、実際の評価項目はそこに留まらず、組織を全体的に見渡すものとなり、財務関係や法令遵守なども見てもらった点が良かった。
	（ウ）データ収集源・収集方法の適切性	・ アンケートや聞き取り調査は精力的に実施してもらい、特にアンケートは協議を重ねて何回か作り直してもらった。その結果、アンケート回答者については意識・意見などある程度わかったのが良かった。また、回答者が少なかった点は残念であり、アンケート未回答者の意識や意見も知りたいところであった。
評価の有用性 = 評価結果は団体にとって有益であったか。それはなぜか。	（ア）教訓・提言などの受け入れ団体にとっての有用性	・ 指摘された内容は「おっしゃるとおり」という感じであった。運営メンバーにとっては「分かっているが手が回らない」点でもあるが、再認識した部分が多かった。 ・ ボランティアや支援者に対して評価結果を報告してもらう予定なので組織全体でこれら問題を共有し、改善につなげたい。
	（イ）評価報告内容の分かり易さ（報告書の質、発表等）	・ わかりやすいと感じた。
インターン制度の有効性 = インターン制度は有効に機能していたか。それはなぜか。	（ア）インターンの資格、インターンの実施能力の適性度	・ 社会人経験ありの方であったことが良かったように思う。
	（イ）調査実施体制の適性度、インターンの人数、アドバイザーの役割	・ 実施体制およびインターンの人数は、今回程度でよいと思う（アドバイザーに関しては後述あり）。
	（ウ）事前オリエンテーションの内容	・ 評価を受ける側は初めての経験であるため、オリエンテーションの席で「何を評価してもらいたいのか」をアドバイザーを含めて深めていけたのが良かった。 ・ 「何を」「どのように」「何のために」評価する或いはさせるのか、メンバーのもつイメージは当初「ばらばら」だったので、できれば会のメンバーが多数いる場で行えれば、なお良かったのではないと思う。 ・ その後、評価インターンと何度か協議して最終的にまとまったが、評価の目的や評価手法を決めるまでには、時間がかかるものだと感じた。

（出所）「ラオスのこども」による報告をもとに筆者作成

の共有につながる」という評価委託者の意向を反映させた企画であり、評価インターンとしても組織運営の改善にむけて協力できたものとする。

(6) 評価インターン受入団体による評価

日本評価学会評価インターン報告会において、受入団体である「ラオスのこども」より、評価インターンの活動と評価結果に対する評価報告がなされた。これに係わる評価設問は、研修委員会が事前に作成しており、3つの評価項目によって構成されている。受入団体の評価として口頭報告された内容を表3にまとめている（当該内容については第3節以降で触れることとしたい）。

3. 実用重視評価との接点

以上、評価インターンによる評価調査の一例を概観してきたが、当該事例のように「組織運営に関する全体的評価と改善策の呈示」を目的とする評価調査にあっては、いかに組織運営の改善につながるかが問われていると言える。

然るに、評価結果が対象組織にとって有用なものであったか（意味あるものであったか）検討することは、2つの点において重要であるとする。

第1に、評価初学者の“評価体験”に終始せず、評価依頼者に対して意味ある介入とならなければ、そもそも評価目的を満たしたことになること。

第2に、評価インターンによる評価活動が“有用である”と認知されることにより、今後、評価インターン制度の需給が拡大し、中・長期的には評価文化の浸透に貢献しうると考えられること。

そこで本節では、評価インターンによる評価調査について、有用性の視点から評価を試みることにする。その際、評価の有用性に注目した評価方法として、Patton（1997）によって提起された実用重視評価の考え方を援用し、以下に掲げる実用重視評価の6つの視点を手掛かりに考察する。

(1) 実用重視評価の基本的な考え方

実用重視評価は、従来の定量的手法と定性的手法の相克を評価の有用性という尺度で統合しようと、Patton（1997）によって提起された理論であり、本学会においても長尾（2003）がその基本的な考え方や実践事例、問題点について紹介・論考している。長尾によれば、実用重視評価はその基本的な考え方として6つの視点から考察することができる。

- (イ) 評価者の役割
- (ロ) 利害関係者の評価活動への参加
- (ハ) 成果重視の評価デザイン
- (ニ) データの分析、解釈、判定
- (ホ) 提言と報告
- (ヘ) 政治的圧力への対応

実用重視評価で最も重要な原則は、個々の評価の目的に沿って結果の有用性を極大化することであるとされる。有用性の大小は評価結果の活用によって決まるため、上記のいずれの視点においても、その要諦は「結果の活用」に資するか否かにある。例えば（イ）評価者の役割については、評価者はまず評価の目的を見極めるとともに、評価結果の活用に関心をもつと想定される層を特定化することが求められ、また評価の目的別に利害関係者の範囲を考慮して、評価への関与の度合いをコントロールすることが必要とされる（長尾2003, p.60）。

また（ロ）利害関係者の評価活動への参加では、評価結果の想定利用者を評価作業に参加させることにより、関係者相互の事業に関する理解の共有、事業に対する当事者意識の向上、事業への関与の拡大を導き、結果として評価結果の活用を担保できると考える（長尾2003, p.60）。

さらに（ハ）成果重視の評価デザインの視点については、評価の「成果」を評価結果の利用者が望ましいと考える形で規定し、それに基づいて評価を組み立てることにより、結果の活用の確率は高まると考える。したがって評価者は、事業関係者および利害関係者と評価枠組（事業成果の定義、成果指標、データ収集方法）について、繰り返し協議することが重要となる（長尾2003,

p60)。

加えて(二)データの分析・解釈・判定に関しては、通常の評価と基本的には同様とするが、その過程に評価結果の想定利用者を参加させることで、結果に対する当事者意識を持つようになると考える(長尾2003, pp.60-61)。

そして(ホ)提言と報告は、実用重視評価が最も重視するところであり、評価結果の利用について最大限の努力が必要とされる(長尾2003, p.61)。

最後に(ヘ)政治的圧力への対応では、実用重視評価において評価者は、事業と事業関係者に積極的に関わるため、ともすると政治的な綱引きに巻き込まれる場合がある。その際には評価の想定利用者にとって有益な成果が得られるような評価を実施できるよう政治的ゲームに主体的に関与することもあるとしている(長尾2003, p.61)。

(2) 評価インターンはどこまで有用な評価を行ったのか

前項では長尾の論考を拠り所として、実用重視評価の基本的な考え方について触れた。そして、実用重視評価の最も重要な原則が「個々の評価の目的に沿って結果の有用性を極大化すること」であることを確認した。

然るに、評価インターンが実施した評価調査は、これら実用重視評価の要素をどこまでクリアできたであろうか。ここでは評価インターン活動および評価インターン受入団体による評価について、長尾の掲げる6つの視点をもとに振り返ることとしたい。

視点1：評価者の役割

評価の目的および評価結果の想定利用者の特定については、研修委員会が事前に用意した評価仕様書フォーマットを完成できれば、概ね達成できる段取りになっていた。加えて今回は、組織運営の改善を目的とする形成評価の性格が強く、利害関係者の幅広い参加が必要と考えられた。

しかし先述の通り、オリエンテーションで評価仕様書(案)を作成した後、評価委託者とインターンの間で数度にわたる話し合いが必要であったのは、評価目的ひとつとっても、評価委託者の内

部が一枚岩ではなく、各個人の置かれている立場によって評価に対するイメージや期待が様でないことが原因であった。今回のケースではFGIを実施して、これらの共有を図るとともに、運営メンバーの日頃の問題意識を汲んだ評価仕様書を作成することができたことから、実用重視の視点をふまえた評価仕様書になったものと考ええる。

ただし、受入団体評価では、評価アドバイザー同席のオリエンテーションの段階で、幅広い層に参加させて評価仕様書を作成することが提案されているが、今回の評価インターンが早い段階でこれを行えなかったことに起因する改善提案であると思われる。今後、評価インターンがこれを円滑に行えるよう効果的な手法を検討する必要があるだろう。

視点2：利害関係者の評価活動への参加

視点1と重複するが、今回の評価調査では、評価仕様書の作成段階で運営メンバー(理事および事務局)の協力を得たことがあげられる。また、データ収集の一環として実施されたアンケート調査では、質問内容の構成について運営メンバーと3度の協議・修正を行って評価委託者の意向、すなわち「知りたいこと」を洩らさないように設計しており、同調査に対する関心・期待を高めることとなった。

その結果、受入団体評価では、アンケート作成過程及びアンケート回答結果について参考になったと評価されていることから、実用重視評価の要素に対応できたものと考ええる。

ただし、利害関係者と言っても運営メンバーの他にボランティアや資金提供者がおり、これら主体の積極的な参加を得ているわけではないため、今後、組織評価であっても評価活動への参加の度合いを高める方策を検討する必要があるだろう。

視点3：成果重視の評価デザイン

ここでいう成果重視の意図は、ロジカル・フレームワークやプログラム・セオリー・マトリックスを用いて、事業における「投入・活動・結果・成果」の因果関係を明示しつつ、ターゲットグループにとって望ましい事業成果を定義して評価デザインを行うことにあるが、今回のケースは組織

運営の改善を目的とする組織評価であるため、視点3とは直接的には関連性が薄い。

視点4：データの分析・解釈・判定

今回の評価調査では、JANIC基準を用いてレーティングを実施しているが、数値が出揃った段階で評価委託者に対して経過報告を行った。先述の通り評価委託者は同基準を経験済であったから、不明瞭な箇所や自己認識と大きく異なる数値（特に認識より低い数値）について関心が集中した。評価インターンは根拠を示して説明を行うことで納得を得たが、なかにはインターン側の認識不足や事実誤認により、再レーティングを行った箇所もあった。

また、受入団体評価においては、組織を全体的に見渡した点および財務関係・法令遵守についても判断根拠を示してレーティングした点も評価されており、結果として実用重視の視点を加味できたものと思われる。

ただし、今回のケースのように評価委託者にとって既知の基準を用いてレーティングを行うのはむしろ稀であり、一般的には評価インターンにとって、データ分析・解釈・判定の過程に実用重視評価の要素を取り入れるには、より周到な準備が必要であると思われる。

視点5：提言と報告

提言と報告については実用重視評価が最も重視するところであるが、提言・改善策を含む評価結果概要について「ラオスのこども」定期総会¹¹の場において報告した。

これは「より多くの関係者が集まる定期総会の場で評価結果報告会を実施することで、組織の問題点の共有につながり、組織にとってプラスになる」という森共同代表の考えを反映させたものであった。当日、会場からは「組織のもつ強み弱みが分かった」「評価報告が分かりやすく楽しく聞けた」とのコメントが寄せられた。運営メンバーだけでなく、組織全体で問題点の共有化を図れるという意味で、定期総会等における評価調査報告は有効であると考えられる。

受入団体評価では「運営メンバーにとって“分かっているが手が回らない”点が指摘され、課題

を再認識した部分が多かった」との評価が寄せられた。その意味で提言・改善策は評価委託者にとって想定範囲内であったようである。ただし、評価結果報告書の提出が平成17年3月、定期総会が同年8月であったが、この間に提言・改善策のいくつかの項目については、すでに改善の取り組みを始めており、少なからず組織運営改善の契機になったものと考えられる。

視点6：政治的圧力への対応

今回のケースについては、組織評価であったこともあり、事業や事業関係者との政治的交渉に関与する要素はなかった。

4．まとめ - 評価インターン制度の課題と制度拡充に向けた提言 -

以上、実用重視評価の基本的視座より、評価インターンの行った組織評価の事例を振り返りながら、その有意性を眺めてきた。その結果、評価インターンの評価事例にも、実用重視評価の要素が認められる箇所があったといえる。例えば「評価者の役割」「利害関係者の評価活動への参加」「データの分析・解釈・判定」「提言と報告」の各視点では不十分ながらも実用重視評価を実践できたと考える。しかしながら、同時に実用重視評価導入の難しさや評価インターンに関わる問題点も見出されることとなった。

そこで以下では、評価インターン制度の課題と今後の制度拡充に向けた提言について4点指摘したい。

第1に実用重視評価の問題点のひとつとして「倫理的問題」があることについては、長尾（2003, p.64）も指摘しており、評価結果の想定利用者の重視する姿勢は、腐敗の危険性を招きかねない点がある、とされる。これに対しては、評価者が特定グループの利益や関心だけでなく、一般的・社会的利益への配慮とのバランスをいかにとるかが重要と指摘されている。したがって、実用重視評価の実践には、高度な倫理性を要求されるが、そもそも評価インターンでは、評価者倫理に

についてはオリエンテーションのなかでも特段に触れられていない¹²。

奇しくも受入団体評価の「インターンの資格」について、「社会人経験ありの方であったことが良かった」とする評価コメントが寄せられたのは、評価者倫理の重要性と関係があると思われる。評価インターンに何らかの倫理性規準を遵守させることも一案である。

第2に評価インターンは、評価委託者との協議のなかから評価仕様書を作成するのに思いのほか時間を取られるが、仮に幅広く参加者を募って評価仕様書を作成しようとする、その効果的な方法論をいまのところ確保できていない。恐らく通常の会議形式では非効率であるし、FGIでは全員参加は困難である。

確かに、参加型ミニワークショップを開催することで、参加者に共通イメージの構築とその後の評価活動への参加が期待できるが、ワークショップ運営にも特殊の技能を要することから、果たして評価インターンにできるものか新たな課題が生じる。評価仕様書の効果的な作成方法は、したがって早急に検討すべき課題である。

第3に、評価結果の報告方法については、基本的に評価委託者の考え方によるところが大きく、評価報告書の提出で完了するケースもあれば、評価報告会を開催するケースもあるだろう。

今回の評価インターンの経験からすれば、組織評価にあっては組織運営幹部のみならず、関係者全体に向けた報告会の開催が望ましいものと考えられる。評価インターンと評価委託者のマッチングの際に、評価報告会の開催について推奨するなどの措置も一考に値する。

第4に、実用重視評価の問題点のひとつに「評価者の質の問題」が指摘され、実用重視評価の実践には評価者の知識、経験、能力が大きく依存する（長尾2003, pp.64-65）とされる。しかも、実用重視の評価者に求められる交渉者の役割は、評価実務を初めて経験するインターン単独ではそもそも困難であり、評価アドバイザーや学会研修委員会のバックアップが不可欠であると言える。したがって、現行のアドバイザー体制をさらに拡充させ、評価者の質の問題を回避できるようなバックアップ体制の検討についても、検討する価値は

あると言えるだろう。

最後に、評価インターン制度の一層の拡充の必要性を指摘したい。当該制度の目的とするところは、大きく「評価能力の向上」と「評価文化の浸透」があるが、筆者の評価インターンの参加経験から言えば、仕様書の検討から報告書の作成に至るまで、時間的・費用的制約のなかで実施できたことは、評価初学者にとって評価能力の向上に直結する経験であったと考える。16年度の評価インターンに参加した学生からも「当初の評価計画段階で苦労したが、最終的には改善策について数多く言及できた」或いは「評価依頼者とやり取りしつつ、アンケート質問項目の修正・作成した過程で学ぶことが多かった」、さらには「評価することの難しさと、評価されたものをさらに読みとくことの難しさを体感した」との感想が寄せられている。これらの言葉が示すように、評価インターン制度は初学者ゆえの試行錯誤はあるものの、参加者の評価能力向上に有効であると考えられる。今後、一層の制度拡充によって、評価研究の裾野が広がっていくことと思われる。

謝辞

本稿で紹介する評価事例は、平成16年度日本評価学会評価インターン出前サービス事業において山田孝子、山村真紀両氏と評価チームとして取り組んだ成果である。評価インターンにおいては「ラオスのこども」森透共同代表はじめ、会の皆様に大変お世話になった。

また、本稿執筆にあたり立教大学大学院21世紀社会デザイン研究科の入山映教授および査読者の方々から、多くの示唆をいただいた。ここに記して深く御礼申し上げたい。

注記

- 1 例えば西出（2003）の紹介事例など。その他、全米評価学会の公式ホームページでは、評価研究で修士号・博士号を付与する高等教育機関を紹介している。詳細は下記を参照。

http://www.eval.org/university_programs.htm

- 2 本件は、平成16年度日本評価学会評価インターン出前サービス事業として実施されたものであり、詳細は当該調査報告書のほか、茂木・山田・山村(2004)による。
- 3 活動内容等の詳細は公式ホームページを参照。
<http://www.deknoylao.org>
- 4 <http://www.hungerfree.net/hyouka.html>
- 5 <http://www.tvac.or.jp/infolib/runners/sm-eval.html>
(2004/12/5当時、現在はリンク切れ)
- 6 田中弥生「NGOによるNGOの免税認証機関研修報告書」、笹川平和財団、1999年
- 7 <http://www.janic.org/janic/ngonet/acseminer.html>
- 8 特定非営利活動法人シェア(国際保険協力市民の会)事務局長の山口誠史氏より「JANIC基準は、もともとNGO関係者に簡便に内部評価を実施してもらえよう設計されているため、使いやすい」との参考意見を得た(2004/11/12聞き取り)。
- 9 なお、本調査では組織運営の評価に特化した活動を行い、同団体が実施するプロジェクト評価は行わないものとして評価仕様書を作成しているため、「受益者に対するインタビュー」を含めていない。
- 10 トライアングレーション(三角検証)については国際協力機構(2004, p.91)を参照。
- 11 平成17年8月13日、東京都大田区内で開催された。
- 12 評価者の倫理性については、評価インターン報告会の際、評価アドバイザーより全米評価学会の定める評価者基準原則について簡単に紹介があった。全米評価学会は' Guiding Principles for Evaluators 'として以下の5つの原則を示している。

Systematic Inquiry

Competence
Integrity/Honest
Respect for People
Responsibilities for General and Public Welfare
詳細は全米評価学会公式ホームページを参照。
<http://www.eval.org/Guiding%20Principles.htm>

参考文献

- 国際協力機構(2004)『プロジェクト評価の実践的手法』、国際協力出版会、p.39
- 長尾眞文(2003)「実用重視評価の理論と課題」、『日本評価研究』3(2): 57-69
- (2005)「評価人材育成: 日本評価学会のアプローチ」、『日本評価学会第6回全国大会発表要旨集録』、pp.149-150
- 西出優子(2003)「市民参加型のNPO評価 - 米国の事例より - 」、『日本評価学会第4回全国大会発表要旨集録』 pp.151-156
- 日本評価学会(2005)「研修委員会2004年度活動報告」、『日本評価研究』5(1): 57
- 茂木勇・山田孝子・山村真紀(2004)「ラオスのこどもの組織運営に関する評価調査」、『日本評価学会第5回全国大会発表要旨集録』、p.277
- Patton, M. (1997). *Utilization-Focused Evaluation: The New Century Text*. 3rd Edition. Thousand Oaks, CA: Sage Publications. 日本語抄訳版: マイケル・クイン・パットン著(大森彌監修、山本泰・長尾眞文編)(2001)『実用重視の事業評価入門』、清水弘文堂書房(2006.1.30受理)

A Study of Student Intern Program for Its Promotion: The Case of Evaluation on Nonprofit Organization Operation of “Action with Lao Children” (DeknoyLao)

Isamu Moteki

Graduate School of Social Design Studies, Rikkyo University
isamsan136@ybb.ne.jp

Abstract

Evaluator training systems in Japan have just started and "The Japan Evaluation Society (JES)" gives great impacts on their promotion, especially student intern program managed by training committee seems to be successful as its dawn.

Therefore, it can safely be said that student intern program mainly aims at evaluator training, but it doesn't mean there's no need to pay attention to the quality of evaluation. If student intern neglects "utilization of evaluation findings" at all, time and cost participants(clients) paid will be wasted because neither reports nor comments are useless for most of them.

This paper first outlines the case study of student intern's evaluation on Japanese nonprofit organization and then examines the utility of it in the context of utilization-focused evaluation. Consequently, looking at the case with six viewpoints of utilization-focused evaluation, some positive elements and challenges are found out through it. In order to develop our evaluator training systems hereafter, student intern should pay more attention to utilization of evaluation findings based on utilization-focused evaluation..

Keywords

utilization-focused evaluation, utilization of evaluation findings,
organization evaluation, evaluator training, student intern

日本評価学会第6回全国大会
開催報告

1. 第6回全国大会プログラム(実績)

2. 共通論題セッション報告

共通論題	「行政評価と行政の質の保証 外部評価を中心に」	窪田 好男(神戸学院大学)
共通論題	「評価人材の育成-大学院レベルの評価教育への期待」	長尾 眞文(広島大学)
共通論題	「公益法人制度と評価の役割」	田中 弥生(東京大学)
共通論題	「セクター評価によるODAの質の向上 (JICA総合的な評価:総合分析を題材として)」	三好 皓一(立命館アジア太平洋大学) 渡辺 博(東洋エンジニアリング)
共通論題	「評価の評価-評価の質保証のための仕組みと手法」	牟田 博光(東京工業大学)
共通論題	「マニフェストと評価」	岩淵 公二(政策21) 吉川 富夫(県立広島大学)
共通論題	「環境評価と政策評価」	松岡 俊二(広島大学)

3. 自由論題セッション報告

自由論題	「行政評価ひろしまセッション」	吉川 富夫(県立広島大学)
自由論題	「環境協力評価」	梅田 次郎(日本能率協会コンサルティング)
自由論題	「ODA評価の課題」	長谷川 弘(広島修道大学)
自由論題	「政策評価の課題」	三好 皓一(立命館アジア太平洋大学)
自由論題	「ODA評価手法」	山谷 清志(同志社大学)
自由論題	「保健衛生評価」	石田 洋子(コーエイ総合研究所)
自由論題	「自治体による評価の課題」	梅内 拓生(国際医療福祉大学)
自由論題	「開発能力アセスメント」	小野 達也(鳥取大学)
自由論題	「教育評価」	松岡 俊二(広島大学)
自由論題	「日本評価学会最初の5年の歩み」	齋藤 貴浩(大学評価・学位授与機構)
自由論題	「満足度意識調査の評価活用」	古川 俊一(筑波大学)
自由論題	「ジェンダー評価」	星野 敦子(十文字学園女子大学)
自由論題	「評価インターンによる調査報告」	村松 安子(東京女子大学)
		源 由理子(国際開発コンサルタント)

日本評価学会第6回全国大会
開催のご報告とお礼

2005年12月10-11日開催(広島大学)の日本評価学会第6回全国大会には180名を超える多くの方にご出席いただき誠にありがとうございました。会員各位の日頃の研究や実践活動の報告を基に、評価研究者、実務者の間の経験、情報、知識を共有化する場として、皆様にとって有益な機会となりましたら幸甚に存じます。至らない点多々あったかと存じますが、多くの方のご尽力により本大会を開催することができましたこと、心より感謝申し上げます。今後とも当学会の活動にご高配賜りますようどうぞ宜しくお願い申し上げます。

第6回全国大会

実行委員会委員長

プログラム委員会委員長

石井 眞治

長尾 眞文

日本評価学会第6回全国大会「評価による質の保証」(於：広島大学)
プログラム(実績)

第1日：2005年12月10日(土)

9:00-9:45	受付			
午前の部 9:45-11:45	自由論題 行政評価ひろしま セッション (吉川富夫、梅田次郎)	自由論題 環境協力評価 (長谷川弘)	自由論題 ODA評価の課題 (三好皓一)	自由論題 政策評価の課題 (山谷清志)
お昼休み 11:45-13:00	お昼休み/ポスターセッション (理事会11:50~12:50)			
午後の部 13:00-15:00	共通論題 行政評価と行政の質の保証 外部評価を中心に (窪田好男)	共通論題 評価人材の育成 大学院 レベルの評価教育への期待 (長尾真文)	自由論題 ODA評価手法 (石田洋子)	
午後の部 15:15-17:15	共通論題 公益法人制度と評価の役割 (田中弥生)	共通論題 セクター評価によるODAの質の 向上(JICA総合的な評価:総合分 析を題材として) (三好皓一、渡辺博)	自由論題 保健衛生評価 (梅内拓生)	
17:30-18:30	総会			
19:00-20:30	懇親会 広島大学附属中・高等学校・食堂			

第2日：2005年12月11日(日)

9:00-9:30	受付			
午前の部 9:30-11:30	共通論題 評価の評価 評価の質保証 のための仕組みと手法 (牟田博光)	自由論題 自治体による評価の課題 (小野達也)	自由論題 開発能力アセスメント (松岡俊二)	
お昼休み 11:30-12:45	お昼休み/ポスターセッション			
午後の部 12:45-14:45	共通論題 マニフェストと評価 (岩淵公二、吉川富夫)	共通論題 環境評価と政策評価 (松岡俊二)	自由論題 教育評価 (齋藤貴浩)	
午後の部 15:00-17:00	自由論題 日本評価学会 最初の5年の歩み (古川俊一)	自由論題XI 満足度意識調査の 評価活用 (星野敦子)	自由論題XII ジェンダー評価 (村松安子)	自由論題XIII 評価インターンによる 調査報告 (源由理子)

()内は座長

自由論題報告の時間は原則として、発表20分、質疑8分、交替2分(計30分)です。

前日企画：2005年12月9日(金)

学校評価シンポジウム：「学校評価システムによる義務教育の質の向上」

14:00~16:30 広島大学東千田キャンパス大講義室

特別研修会：「参加型評価の理論と実践 協働社会の推進をめざして」

10:00~17:00 広島大学東千田キャンパス第3演習室

12月10日(土)

午前の部(9:45-11:45)

自由論題セッション 行政評価ひろしまセッション
座長 吉川富夫(県立広島大学) 梅田次郎(日本能率協会コンサルティング)

9:45-10:13
広島市の行政評価制度について
糸山隆(広島市)

10:15-10:43
三次市の行政評価システム「The 行政チェック」
奥野通(三次市)

10:45-11:13
自治体の行政評価と人事評価 - 今後の自治体改革に向けて -
佐藤幹(ローカル・マニフェスト推進ネットワーク中国)

11:15-11:43
市町村財政情報をわかりやすく示す工夫
~住民が財政を語るために~
平野誠治(広島県)

自由論題セッション 環境協力評価
座長 長谷川弘(広島修道大学)

9:45-10:13
環境国際協力評価活動に関する一考察と提言: わが国と欧米
諸国と国際機関との比較
廣野良吉(政策研究大学院)

10:15-10:43
北東アジアにおける経済発展段階に応じた地域環境協力の形成
青正澄(共立総合研究所)

10:45-11:13
持続可能な開発評価の基礎的研究(2)環境税導入国を中心とし
た経済・税・環境との関係の検討
泉田晃彦(国際医療福祉大学大学院・院生)
梅内拓生(国際医療福祉大学大学院)

11:15-11:43
廃棄物分野における評価活動
和田英樹(廃棄物政策研究所)

自由論題セッション ODA評価の課題
座長 三好皓一(立命館アジア太平洋大学)
討論者 松本哲男(名古屋大学)

9:45-10:13
援助の「妥当性」の評価に関する考察
三輪徳子(国際協力機構)

10:15-10:43
開発援助事業におけるプロセス評価に関する考察
~コミュニティ開発型プロジェクトを事例として~
源由理子(国際開発コンサルタント)

10:45-11:13
費用便益分析を用いたODA水供給プロジェクトの効率性評価
に関する研究

永吉洋之(東京大学大学院・院生)
山田淳(立命館大学)
Victor Muhandiki(立命館大学)
佐伯健(国際協力機構)

11:15-11:43
合同評価を通じた評価体制整備への取組
野田邦雄、速見公子、鈴木いづみ、
嶋谷哲(国際協力銀行)

自由論題セッション 政策評価の課題
座長 山谷清志(同志社大学)

9:45-10:13
公的研究機関の技術評価
原田隆(産業技術総合研究所)

10:15-10:43
政策評価にみる本質的矛盾 アカウンタビリティと経営
古川俊一(筑波大学)

10:45-11:13
途上国の財政制度と財政執行の評価方法: PEFA プログラム
の業績測定枠組みの考察
上野宏(南山大学)

11:15-11:43
日本の構造調整借款20年間のレビューからみる日本政府の政
策判断
田中弥生(東京大学)

午後の部(13:00-15:00)

共通論題セッション 行政評価と行政の質の保証
外部評価を中心に
座長 窪田好男(神戸学院大学)
討論者 梅田次郎(日本能率協会コンサルティング)

1. 評価と民主的コントロール 「信頼できる第三者」による
評価の可能性
砂原庸介(東京大学大学院・院生)
2. 政策評価と政策研究 岩手県「県民参加型外部評価」の
取組み
南島和久(岩手県立大学)
3. 独立行政法人評価の展開と今後の課題 通則法枠外の評価
環境をめぐる
西山慶司(法政大学大学院)

共通論題セッション 評価人材の育成
大学院レベルの評価教育への期待
座長 長尾眞文（広島大学）

1. 大学院レベルにおける評価教育・問題提起
入山映（立教大学大学院）
2. 教育学研究科における評価教育～教職大学院を中心として～
中原忠男（広島大学大学院）
3. 大学評価の現状からの問題提起
米澤彰純（大学評価・学位授与機構）
4. 大学院レベルの評価人材教育への期待～大学院生の視点より～
茂木勇（立教大学大学院・院生）

自由論題セッション ODA評価手法
座長 石田洋子（コーエイ総合研究所）

13：00-13：28

理論に基づいた開発援助の評価
山下道子（国際協力銀行）

13：30-13：58

開発事業におけるインパクト評価手法改善への取組み
小林信行、魚住麻子、野田邦雄（国際協力銀行）

14：00-14：28

ベキ法則に基づく健康と経済のマクロ指標を用いた政府開発援助政策評価法の基礎研究（IV）
正木朋也（統合社会科学研究会）
梅内拓生（国際医療福祉大学）

14：30-14：58

ODA国別評価手法についての考察
駒澤牧子（アースアンドヒューマンコーポレーション）
金子真知（アースアンドヒューマンコーポレーション）

午後の部（15：15-17：15）

共通論題セッション 公益法人制度と評価の役割
座長 田中弥生（東京大学）

パネリスト

横田信孝（内閣官房）

竹内正興（国際開発センター）

工藤泰志（言論NPO）

古川俊一（筑波大学）

共通論題セッション セクター評価によるODAの質の向上
（JICA総合的な評価：総合分析を題材として）
座長 三好皓一（立命館アジア太平洋大学）、渡辺博（東洋エンジニアリング）

1. 総合的な評価の必要性について
三好皓一（立命館アジア太平洋大学）、
渡辺博（東洋エンジニアリング）
2. JICAにおける「評価結果の総合分析」の経験と教訓
三輪徳子（国際協力機構）
3. 「評価結果の総合分析・初中等教育/理数科分野」について
浜野隆（お茶の水女子大学）
4. 評価結果の総合分析・情報通信技術/IT人材育成と各分野におけるIT活用
大嶋淳俊（UFJ総合研究所）

自由論題セッション 保健衛生評価
座長 梅内拓生（国際医療福祉大学）

15：15-15：43

わが国の健康づくり指標と評価
健康日本21とローカル・マニフェスト
伊関友伸（城西大学）

15：45-16：13

保健医療薬剤政策評価の基礎研究（2）-循環器系薬剤を中心に-
高他武始（国際医療福祉大学大学院・院生）
梅内拓生（国際医療福祉大学大学院）

16：15-16：43

適切な医薬品使用に関する考察（2）-ストレス関連疾患を取り上げて-
梅原貞臣（国際医療福祉大学大学院・院生）
梅内拓生（国際医療福祉大学大学院）

16：45-17：13

保健医療プロジェクトの新しい学際的な評価法に関してのケーススタディ
仲佐保（国立国際医療センター）
中村安秀（大阪大学）、草郷孝好（大阪大学）
池田光穂（大阪大学）、天沼直子（大阪大学・院生）
倉沢愛子（慶應大学）

総会（17：30-18：30）

懇親会（19：00-20：30）

12月11日(日)

午前の部(9:30-11:30)

共通論題セッション 評価の評価
評価の質保証のための仕組みと手法
座長 牟田博光(東京工業大学)

1. 外部評価委員会による評価の評価 - 総務省政策評価・独法評価委員会を例に -
山谷清志(同志社大学)
2. 評価方式による行政監査の取組について
井畑晃洋(三重県)
3. 外部有識者によるODA評価の二次評価
牟田博光(東京工業大学)
4. 評価の評価: 大学評価・学位授与機構による試行的大学評価の検証方法とその結果
齋藤貴浩(大学評価・学位授与機構)

自由論題セッション 自治体による評価の課題
座長 小野達也(鳥取大学)

- 9:30-9:58
基礎自治体における参加型評価に関する一考察～相模原市村富公園再整備事業を事例として～
茂木勇(立教大学大学院・院生)
- 10:00-10:28
公園を介した官民連携
山崎肇(県立広島大学大学院・院生)
- 10:30-10:58
問題指摘着目型の法制度の実態評価～その類型と試行: リサイクル法制度を例に～
田崎智宏(国立環境研究所)
- 11:00-11:28
自治体評価にみる本質的課題とその解決方法に関する考察
田中啓(静岡文化芸術大学)

自由論題セッション 開発能力アセスメント
座長 松岡俊二(広島大学)

- 9:30-9:58
社会的能力アセスメントに基づく援助投入評価
高橋与志(広島大学) 松岡俊二(広島大学)
工藤高志(広島大学・院生)
- 10:00-10:28
後発移行経済国タジキスタン政府の開発マネージメント・キャパシティ・アセスメント
登丸求己(玉川大学)
- 10:30-10:58
環境対策における社会的能力の評価
村上一真(広島大学・院生) 松岡俊二(広島大学)

午後の部(12:45-14:45)

共通論題セッション マニフェストと評価
座長 岩淵公二(政策21) 吉川富夫(県立広島大学)
討論者 山谷清志(同志社大学)

1. 問題提起
吉川富夫(県立広島大学)
2. わが国におけるローカル・マニフェストと評価の現状と課題
西尾真治(UFJ総合研究所)
3. マニフェスト主導型の行政運営に向けた取組み
鎌田徳幸(岩手県)
4. マニフェストとアカウンタビリティ
平松英哉(同志社大学大学院)

共通論題セッション 環境評価と政策評価
座長 松岡俊二(広島大学)

1. 経済学からみた環境評価とその政策利用
竹内憲司(神戸大学)
2. 生態学からみた環境評価とその政策利用
中越信和(広島大学大学院)
3. 開発援助における環境社会配慮
渡辺泰介(国際協力機構)
4. 開発プロジェクトにおける環境影響評価と経済評価の統合
長谷川弘(広島修道大学)

自由論題セッション 教育評価
座長 齋藤貴浩(大学評価・学位授与機構)

- 12:45-13:13
数学的リテラシー(PISA)に影響を及ぼす国別要因の分析
渡邊美紀(東京工業大学大学院・院生)
牟田博光(東京工業大学大学院)
- 13:15-13:43
Effectiveness of Piloting Activities on Students' Attitude
(Case Study of IMSTEP, Bandung, West Java)
Harun, Imansyah(Indonesia University of Education),
Ogasawara, Aimi (Tokyo Institute of Technology),
Muta, Hiromitsu (Tokyo Institute of Technology)
- 13:45-14:13
学校評価の実践: 広島市の公立校4校の事例を通して
田畑直子(メルボルン大学・院生)
- 14:15-14:43
人間性教育と性教育評価の基礎 第2報 ~社会の動向、性行動の動態と教育政策の検討~
高田恵子(国際医療福祉大学大学院・院生)
梅内拓生(国際医療福祉大学大学院)

午後の部 (15:00-17:00)

自由論題セッション 日本評価学会最初の5年の歩み
座長 古川俊一(筑波大学)

15:00-15:28

ロッシら「プログラム評価：システマティックアプローチ第7版」の改訂プロセスに見る米国 プログラム評価研究の発展と日本における課題
大島巖(東京大学大学院)

15:30-15:58

日本評価学会5年の歩みと政策評価、ODA評価
山谷清志(同志社大学)

16:00-16:28

自治体行政評価の動向 2001年から2005年まで
小野達也(鳥取大学)

自由論題セッション 満足意識調査の評価活用
座長 星野敦子(十文字学園女子大学)

15:00-15:28

大学生の意識調査をベースとした日本のODAの説明責任に対する間接評価
高千穂安長(玉川大学)

15:30-15:58

韓国におけるサイバー大学生の満足度に影響する要因分析及び学習環境把握
金性希(東京工業大学大学院・院生)
牟田博光(東京工業大学大学院)

16:00-16:28

自治体公共サービス満足度ランキング向上の方法論
大島章嘉(市民満足学会/ワード研究所)

16:30-16:58

大学における授業改善をはかる要因類型化手法
星野敦子(十文字学園女子大学)
北原俊一(十文字学園女子大学)
新行内康慈(十文字学園女子大学)
安達一寿(十文字学園女子大学)
綿井雅康(十文字学園女子大学)
牟田博光(東京工業大学大学院)
コメンテーター 齋藤貴浩(大学評価・学位授与機構)

自由論題セッション ジェンダー評価
座長 村松安子(東京女子大学)
討論者 大沢真理(東京大学)

15:00-15:28

日本学術会議におけるジェンダー主流化への胎動
原ひろ子(城西国際大学大学院)

15:30-15:58

ジェンダー関係の平等化推進と成人学習の評価
國信潤子(愛知淑徳大学大学院)

16:00-16:28

北九州市男女共同参画基本計画進捗状況評価
三隅佳子(アジア女性交流・研究フォーラム)
山田貴広(北九州市役所)

16:30-16:58

男女共同参画の視点から見た環境施策 北九州市の事例を手がかりとして
織田由紀子(アジア女性交流・研究フォーラム)

自由論題セッション 評価インターンによる調査報告
座長 源由理子(国際開発コンサルタント)

15:00-15:28

広島県立教育センターの事業改善評価
<派遣先:広島県立教育センター>
原口裕紀子(立教大学大学院・院生)
山田孝子(立教大学大学院・院生)
コメンテーター:寺地靖仁(尾道市教育委員会)

15:30-15:58

学校評価表の開発とその課題<派遣先:尾道市栗原北小学校>
串本剛(広島大学大学院・院生)
渡辺明美(早稲田大学大学院・院生)
コメンテーター:森田哲(広島県立広島中学校)

16:00-16:28

メロウ倶楽部の事業評価 <派遣先:(特活)メロウ倶楽部>
藪並郁子(立教大学大学院・院生)
コメンテーター:西出順郎(琉球大学)
駒澤牧子(アースアンドヒューマンコーポレーション)

16:30-16:58

NPO法人岡山NPOセンター(組織評価)
佐藤浩子(立教大学大学院・院生)
中川力介(早稲田大学大学院・院生)
コメンテーター:西出順郎(琉球大学)

ポスターセッション(12/10, 11お昼休みに開催)

特別セッション:学校評価研修プログラムの認証制度について

1. 認証制度の意義
小室雪野(国際開発センター)
2. 教員評価者研修導入の事前プロセス
寺地靖仁(尾道市教育委員会)
3. 教員評価者研修の実際
池田武彦(広島県立教育センター)
4. ポスターセッション:広島県における学校評価の事例
東広島市立西条小学校 (高木美佐恵)
三次市立清河小学校 (池尻まゆみ)
尾道市立栗原北小学校 (石田謙豪)
広島県立教育センター (斉藤美由紀)
広島県立広島中学校 (森田哲)
広島県立府中東高等学校 (柴田武秀)
広島市立基町高等学校 (菅本和秀)

共通論題セッション報告

共通論題セッション 「行政評価と行政の質の保証 外部評価を中心に」

座長 窪田 好男（神戸学院大学）

討論者 梅田 次郎（日本能率協会コンサルティング）

本セッションでは窪田好男（神戸学院大学）を司会として3つの報告が行われた。いずれの報告についても、討論者の梅田次郎（日本能率協会コンサルティング）との間で議論が行われるとともに、フロアとの討議も活発に行われた。

まず砂原庸介氏（東京大学）の「評価と民主的コントロール 『信頼できる第三者』による評価の可能性」が報告された。本報告では、政治学の業績に依拠しつつ、実際に評価情報を活用する主権者にとってどのような評価情報が説得的かという議論を展開した。評価論に新たな視点を導入する重要な指摘である。他のアプローチで本テーマを深める余地もある、言い換えれば新規参入の余地も大きいと感じた。

ついで南島和久氏（岩手県立大学）の「政策評価と政策研究 岩手県『県民参加型外部評価』の取組み」が報告された。本報告では、NPOの政策21の活動を例に政策研究型の外部評価をあるべき姿として提示した。評価は行政が行うものというイメージに対し、外部主体が自立的・主導的に行うこともできるという事実を提示したことは重要である。

最後に西山慶司氏（法政大学）の「独立行政法人評価の展開と今後の課題 通則法枠外の評価環境をめぐって」が報告された。本報告は、制度開始後4年を経過した独立行政法人評価について、内閣レベルのトップダウン的決定、業績評価の顕著な分散化、内閣の関与の実質的な制度化など、当初は想定されなかった変化が生じていることを論じた。国や自治体の政策評価に比べて注目されにくい独立行政法人評価を取り上げた貴重な業績である。

なお、報告者とフロアとの間で活発な質疑応答が行われたが、それにとどまらず、以後の他のセッションで本セッションの報告内容が繰り返し引用された。この事実は、本セッションのインパクトの大きさを物語るのではないだろうか。

現在、外部評価の主流は学識者・有識者による委員会方式であるが、その有効性等への疑問 とりわけ、本来は政治が果たすべき役割を学識者・有識者の委員会に代行させている が本セッション立ち上げのきっかけの一つとなった。各報告とも、疑問に最終的な解答を提示したとはいえないかもしれないが、それぞれ今後の研究発展の余地の大きいテーマを取り上げており、今後の成果が期待される。

今後も継続して同様のセッションが企画されることを期待して結びとしたい。

共通論題セッション 「評価人材の育成 - 大学院レベルの評価教育への期待」

座長 長尾 真文（広島大学）

本セッションでは、評価慣行の一層の普及・定着にとって、専門性を備えた評価人材の育成が不可欠の課題であるとの認識のもとに、大学院レベルの評価教育的を絞って、異なる立場・視点からその現状と将来的な可能性、および評価学会がその促進に果たし得る役割について討議した。冒頭で入山映氏（立教大学）が「大学院レベルにおける評価教育」について問題提起を行い、中原忠男氏（広島大学）が「(広

島大学)教育学研究科における評価教育～教職大学院を中心として～」、米澤彰純氏(大学評価・学位授与機構)が「大学評価の現状からの問題提起」、茂木勇氏(立教大学大学院)が「大学院レベルの評価人材教育への期待～大学院生の視点より～」のテーマでそれぞれ短い報告をした後、参加者も加わり質疑応答を中心とする活発な意見交換がなされた。

入山氏の問題提起では、「評価」カリキュラムを組む場合の問題点として、学部においてこれに対応する教育を受けた経験のない学生が大部分を占めることが想定され、初学者を対象とするものとならざるを得ないこと、教授陣のavailabilityおよび大学側からする様々な理由双方から、当面開講可能な課目数が限定されること、実務上のスキルとして科目履修を要求あるいは評価するクライアントが存在しないことの3点が挙げられ、同氏の大学院における評価教育経験に基づいて、初学者と実践に必要とされるレベルの間のギャップを埋めるため、理論だけでなく実地の訓練の頻度を増加する必要、その関連での学会の役割の重要性が指摘された。

中原氏は、同氏の大学院教育学研究科の経験を念頭に、従来から教育評価に関わる教育や研究はなされているが、近年における教育界の動向や学校教育の成果に関する社会的関心の高さや説明責任から、目標準拠評価、学校評価、教師評価等に関わる評価方法の開発やそれに基づく評価人材教育といった新たな課題が出ており、平成19年度から開設される予定の「教職大学院」で専門職教育の一環としてとりあげられる見込みと報告。

米澤氏は、大学は公的機関であり、アカウントビリティや質保証・向上の観点から、近年さまざまなレベルで多岐的な評価を行っていること、その関連での確かな自己活動の説明能力や、他者に対して適切で公平な評価を行う能力の重要性が認識されるようになったこと、そして大学院レベルの評価教育のあり方とその実現における連携の可能性について報告。

最後に、茂木氏が、実際に大学院で評価教育を受けた経験から、評価人材教育の議論で、重視されるべき受益者側の視点について報告し、学生の受講満足度も教育効果も評価実習によって高くなること、その意味で昨年度参加した学会の評価インターン・プログラムが極めて有用であったこと、そしてその継続的な実施の可能性を大学と地域の連携による評価人材教育の拡充の方向に認めると述べた。

報告に続く質疑応答・一般討議で一番問題となったのは、評価人材育成の前提として求められる専門性をどのように設定するか、その際に評価専門性と分野別専門性の組み合わせをどう考えるかで、必要とされる評価作業の想定によって大きく異なることが確認された。また、その関連で、専門職大学院教育の枠の中で、評価にどれだけ比重をおくことができるかも議論されたが、現実的な解は「評価副専攻」程度との認識であった。さらに、評価教育で重視されるべき実地演習をいかに実現するかについて、オーストラリア・メルボルン大学の評価業務の受託を教育に活用する経験等の説明もあり、学会評価インターン制度の拡充の可能性についての期待も述べられた。

共通論題セッション 「公益法人制度と評価の役割」

座長 田中 弥生(東京大学)

本セッションでは、115年ぶりに改正される公益法人制度における評価の役割について議論した。討論者は、横田信孝(内閣官房 公益法人制度改革事務局)、竹内正興((財)国際開発センター)、工藤泰志((特)言論NPO)、古川俊一(筑波大学)であり、本制度策定作業に着手中の横田氏より新制度のあらましについて説明があった。

本制度改革の背景には言わば2つの矛盾する目的がある。すなわち、KSD事件などに象徴される公益法人のスキャンダルや天下り問題などの排除と解決という比較的ネガティブな目的、そして「民が担う公」を促進するという積極的な目的である。両者を同時に満たすような制度設計が検討されてきたが、従来

の役所が設立許可をあたえる制度から登記のみで非営利法人が設立できるようになったこと（非営利法人としては登記で設立可能）・登記された非営利法人の中から公益性ある法人が認定されるが、公益性の判断は内閣府のもとにおかれる有識者委員会（地方は知事のもと有識者委員会）が行う、という非営利法人の登記、公益性ある非営利法人認定の2段階構えの方式が採用される。また、公益性が認められた法人に対しては、寄付優遇と収益事業への税優遇の二つが付与されることになる。

この案に対して、二つの目的に照らして本制度が設計されているのか、・すなわち公益性の判断の有無、・公益性を失った法人の整理のあり方が明らかになっているのかとの質問があった。続いて、公益性判断の難しさや政府がそれを行うことへの疑問が投げられた。これに対し、横田氏からは公益性の積極的評価は行わず、公益法人にそぐわない法人をチェックするに留まるだろうとの説明があった。

また、「民が担う公」を促進するには、公益法人自身の自己評価を制度に組み入れる必要があるとの問題が座長より提起された。則ち、政府のチェックは最低限に留め、公益法人自身が評価を行い、それを公開し最終判定は寄付者に委ねるというものである。パネルではこの案を支持する意見が多く、評価方法の一例として、言論NPOが開発したネガティブ・チェック方式が説明された。また、公益法人側の自己評価能力は十分とは言えず、それをサポートするような人材育成、あるいは訓練機会の提供が必要ではないかという意見が出された。新法案制定後は現行2万6千団体の公益法人に加え、新規参入団体の公益性チェック作業が開始されるが、同時に自己評価の視点も寄付者との関係でますます重要視されるだろう。この点について評価学会として何らかの役割を担えるはずであるとの意見が複数のパネルから出された。筆者もその可能性があると思うが、そのためには本制度設計の意味や我が国の民の役割について十分認識をもつ必要がある。

共通論題セッション 「セクター評価によるODAの質の向上（JICA総合的な評価：総合分析を題材として）」

座長 三好 皓一（立命館アジア太平洋大学）
渡辺 博（東洋エンジニアリング）

本セッションではまず、座長から「総合的な評価の必要性について」で、問題提起を行った後、「JICAにおける評価結果の総合分析の経験と教訓」発表者、三輪徳子（国際協力機構）「評価結果の総合分析・初中等教育/理数科分野について」発表者、浜野隆（お茶の水大学）「評価結果の総合分析・情報通信技術/IT人材育成と各分野におけるIT活用について」発表者、大嶋淳俊（UFJ総合研究所）の3つの論題が報告された。

「JICAにおける評価結果の総合分析の経験と教訓」では、特定分野に関する評価結果の総合分析は、総括評価として対象分野における事業結果の全体傾向を示すうえでも、形成評価として当該分野の協力に関する教訓を体系的に抽出するうえでも有用性が高いこと、特に後者は、評価結果のフィードバックの推進に有益であること、既存の評価結果の活用の点で評価コストの節約を可能とすること、1次評価の質が分析の制約要因となり得ること、当該分野や分析の方法論に関する知見など、評価者の総合的な分析能力が分析の質を決める鍵となることが報告された。

「評価結果の総合分析・初中等教育/理数科分野について」では、JICAにより数多く実施された技術協力プロジェクト：初等・中等教育理数科分野について、プロジェクトを整理・類型化し、案件の案件形成・実施過程と協力内容、結果の分析を通じて、貢献要因・阻害要因、今後の案件形成・実施に係る教訓について検討、案件の成否に影響する要因として、特に「計画・立案」「成果の普及手段」「連携」「制度化」「モニタリング・評価」をとりあげ、初等・中等教育（理数科分野）のプロジェクトを展開（計画・実施）する上での課題が報告された。

「評価結果の総合分析・情報通信技術/IT人材育成と各分野におけるIT活用について」では、情報通信技術/IT人材育成に係る対象案件として1997年以降協力を終了した案件7件、IT活用に係る対象案件として教育、保健・医療、貧困、環境の4分野における具体的な活用可能性を示す4分野5案件を選択して実施され、「IT人材育成」については、評価結果の見直し、各案件におけるITの位置づけ、社会状況におけるITの位置づけ、「人材育成」の対象者、協力アプローチ等を基に協力形態の類型化と分析を行い、「各分野におけるIT活用」では、対象分野におけるITの活用形態、IT活用対象者、活用者へのインパクト、ならびに留意点について分析を行ったことが報告され、これらの結果を基に、今後のIT分野における新規案件の計画立案に際して留意すべき教訓や対策を導き出したことが報告された。

共通論題セッション 「評価の評価 評価の質保証のための仕組みと手法」

座長 牟田 博光（東京工業大学）

本セッションの目的は異なるジャンルで行われている各種の評価の評価（二次評価）の取り組みを横断的に取り上げ、評価の品質保証に向けた仕組み、手法のあり方について、議論を深めることであり、以下の4件の発表の後、質疑応答があった。

「外部評価委員会による評価の評価 - 総務省政策評価・独法評価委員会を例に -」（山谷清志）では、政策評価と独立行政法人評価の二次評価として、総務省政策評価・独立行政法人評価委員会の評価を取り上げた。政策評価も独立行政法人評価も、まず内部評価から始まり、評価の客観性を保ち「お手盛り評価」を排する仕組み、プロセスが構築されたが、制度設計段階からいくつかの危惧はあった。評価を実際に運用した後に、予測しなかった問題も出てきた。いろいろな機能不全、予測できなかった課題への対応に追われ、それがメタ評価制度そのものの意義を損ないそうな状況を導いている、との報告があった。

「評価方式による行政監査の取り組みについて」（井畑晃洋）では、三重県の総合計画の施策を対象にした行政監査を取り上げた。「みえ行政経営体系」に基づき行われる評価は執行機関の自己評価であり、監査委員の監査（評価）は独立した第三者評価である。総合計画に定める施策ごとに、施策、基本事業及び事務事業を体系的に3E（経済性、効率性、有効性）Value For Money（金額に見合う価値）及びBest Value（サービスの質と改善可能性の評価）の視点から監査（評価）を実施し、県民に情報を提供しているが、アンケート結果から、行政監査の結果が業務の改善に役立てられている、との回答が多かった、との報告があった。

「外部有識者によるODA評価の二次評価」（牟田博光）では、JICAの事後評価の二次評価を取り上げた。2003年度の評価と比較して2004年度の評価の質が大きく向上した。その理由としては、2003年度の評価結果を基に、2004年2月に「事業評価ガイドライン」の大幅な改定があり、そのガイドラインに沿って一次評価の質の改善が図られた事が大きいと考えられ、二次評価結果の具体的なフィードバックが一次評価の質の向上に効果があった事を示している、との報告があった。

「評価の評価：大学評価・学位授与機構による試行的大学評価の検証方法とその結果」（齊藤貴浩）では、大学評価・学位授与機構が2000年度から2003年度まで国立大学及び大学共同利用機関を主な対象に行った3回の試行的評価結果を取り上げた。「大学の教育研究活動等の改善に役立てる」目的については、ある程度の貢献が認められるものの、大学の個性の促進については、まだ改善すべき点も多く残されている、との報告があった。

共通論題セッション 「マニフェストと評価」

座長 岩淵 公二（政策21）

吉川 富夫（県立広島大学）

討論者 山谷 清志（同志社大学）

本セッションでは、近年のローカル・マニフェストの普及と検証活動が自治体運営に及ぼしている影響と課題を探るべく、本セッションを企画した吉川富夫会員（県立広島大学）から問題提起を行った後、「わが国におけるローカル・マニフェストと評価の現状と課題 発表者、西尾真治会員（UFJ総合研究所）」、「マニフェスト主導型の行政運営に向けた取組み 発表者、鎌田徳幸会員（岩手県）」、「マニフェストとアカウンタビリティ 発表者、平松英哉会員（同志社大学大学院）」の3つの論題が報告された。

問題提起では、ローカル・マニフェストの普及と検証の活動が自治体の運営と評価にどのような影響を与えているのか、行政システムとアカウンタビリティの視点から論点が提起された。それを受け、「わが国におけるローカル・マニフェストと評価の現状と課題」では、わが国の自治体におけるローカル・マニフェストの活用と検証活動の実態を踏まえ、マニフェスト・サイクルの確立に向けた評価のあり方が報告され、マニフェストの質の向上とともに、マニフェストをチェックする住民および議会の関与が提起された。「マニフェスト主導型の行政運営に向けた取組み」では、わが国で最初にローカル・マニフェストを掲げ、マニフェストに基づく県政運営を強力にすすめている岩手県の事例から、マニフェスト主導型の行政運営に必要な予算制度の改革や職員の意識改革の重要性が報告され、マニフェスト導入に伴う行政組織内部のシステム整備のあり方が提起となった。「マニフェストとアカウンタビリティ」では、具体的な政策目標を明記したローカル・マニフェストは首長の責任を明確化することから、自治体運営における政治と行政の責任分担を再構成する可能性が報告され、アカウンタビリティ向上に向けたマニフェストの意義が提起された。

報告後、討論者の山谷清志会員（同志社大学）ならびにフロア会員から報告に対する質問・意見が出され、活発な議論が展開された。報告および質疑を通じ、検証可能性を最大の要件とするローカル・マニフェストの導入が、その機能を十分発揮できていない現状と課題が浮き彫りになるとともに、評価研究の立場からローカル・マニフェストを自治体運営に活かす理論的技術的支援の必要性が実感された。

共通論題セッション 「環境評価と政策評価」

座長 松岡 俊二（広島大学）

本セッションは、生態学（景観生態学）、経済学（環境経済学）などの視点から環境評価が現実の開発政策や環境政策（Policy, Program, Project）評価にどのように使われているのか、また使われていないのかを具体的に議論し、今後の環境評価研究の方向を明らかにする。また、国際開発などの政策実施の現場において、環境評価をどのように利用していくのかについて議論することを目的に設定された。本セッションでは、竹内憲司（神戸大学、環境経済政策学会理事）「経済学からみた環境評価とその政策利用」、中越信和（広島大学、日本景観生態学会会長）「生態学からみた環境評価とその政策利用」、渡辺泰介（国際協力機構）「開発援助における環境社会配慮」、長谷川弘（広島修道大学）「開発プロジェクトにおける環境影響評価と経済評価の統合」の4報告が行われた。

竹内憲司（神戸大学）「経済学からみた環境評価とその政策利用」では、環境の経済評価に関する最新の研究トピックとその政策利用の現状が紹介され、今後の展望が示された。その際、環境の改善や悪化に対する人々の支払い意志額（Willingness to Pay）や受け取り意志額（Willingness to Accept）を、実際の行

動データや仮想的なアンケートへの回答データを用いて推計する手法の動向と問題点が述べられた。

中越信和（広島大学）「生態学からみた環境評価とその政策利用」では、生態学が対象の構造と機能だけでなく、その過程や歴史性、地理的固有性をも加味して研究が行なわれることが紹介された。また、環境評価にあたって、これらの生態学的特性が適用されることは環境の質的側面をより深く評価することを可能にすることが指摘された。

渡辺泰介（国際協力機構）「開発援助における環境社会配慮」では、開発援助における環境社会配慮の特徴をみながら、事業評価と環境社会配慮の関わりについて報告された。環境アセスメントの特徴として、意思決定への参加プロセス、社会的影響、戦略的環境アセスメント、代替案、説明責任と透明性の確保が指摘され、途上国の環境アセスメント制度の特徴やODA評価との関連が述べられた。

長谷川弘（広島修道大学）「開発プロジェクトにおける環境影響評価と経済評価の統合」では、開発途上国における農林業プロジェクトの環境経済評価を対象に、開発プロジェクト評価システムにおける定量的な環境価値の内部化の必要性を踏まえ、従来の環境影響評価と経済評価とを統合する「環境経済評価」の導入について考察するとともに、その効率的実現に向けた評価手法上の課題や実務的重要事項について提案が行われた。

全体討論では、特に、環境の経済的評価手法（費用便益分析）と環境アセスメント手法（公害項目と自然生態系評価）との関係が活発に議論された。大方の意見としては、環境アセスメントと環境の経済的評価は制度的に切り離して行う方が好ましいというものであった。

自由論題セッション報告

自由論題セッション 「行政評価ひろしまセッション」

座長 吉川 富夫（県立広島大学）

梅田 次郎（日本能率協会コンサルティング）

本セッションは、学会開催地に焦点を当てた評価（自治体の行政評価）の課題を集めたというところに特徴がある。広島県の自治体の行政評価は、日本全体の中では先行事例というわけではない。しかしそうであるがゆえに、標準装備化された時代の行政評価、ローカル・マニフェストが先行する行政評価などの新しい課題に直面しているのである。

こうした事例として、「広島市の行政評価制度について」報告者：糸山隆（広島市）「三次市の行政評価システム『The 行政チェック』」報告者：奥野通（三次市）「自治体の行政評価と人事評価 今後の自治体改革に向けて」発表者：報告者（広島市）「市町村の財政情報をわかりやすく示す工夫～住民が財政を語るために～」報告者：平野誠治（広島県）の4つの論題が報告された。

「広島市の行政評価について」では、行政評価導入3年を経た中間的な総括として、ベンチマーク指標と政策との因果関係の把握の困難さ、事務事業レベルでの行政評価の希薄さによる市政改善効果の弱さ、行政評価への住民の関心の薄さなどが、担当セッションの認識として報告された。これに対して、ベンチマーク指標と事務事業指標との混同、行政評価を行う目的の不鮮明さが指摘され、今後は施策中心に行政評価を行ったらどうかという提案があった。

「三次市の行政評価システム『The 行政チェック』」では、行政評価導入1年ではあるが、担当者による一次評価、庁内組織による2次評価、行政チェック市民会議による外部評価という3段階の評価を実施し、予算編成にも反映させている状況が報告された。これに対して、市長マニフェストとの関係に疑問が集中した。市長マニフェストがトップダウンで行政評価の実施を促進した面はあるが、全体としてマニフェストと行政評価の連携はいま試行錯誤中ということがわかった。

「自治体の行政評価と人事評価」については、現在さまざまな自治体で取り込まれている行政評価と人事評価には重複があるので、今後は、行政評価（業績評価）とコンピテンシー評価を分離すべきことが主張された。そして、最近の民間企業でのコンピテンシー評価を取り入れた人事評価の事例、広島市の人事評価の事例などが紹介された。

「市町村の財政情報をわかりやすく示す工夫」では、大規模な合併がすすんだ広島の基礎自治体の財政状況を住民にわかりやすく提示することのできる工夫が紹介された。例えば、自治体の財政を一般的な家庭規模の歳入・歳出、資産・負債にブレイクダウンさせることによって、財政状況が身近な情報に転換されるのである。これに対して、自治体の財政状況は、議員にさえもわかりにくいので、こうした工夫を積み重ねることによって、自治体の自立した経営に貢献できるだろうという支持の声が寄せられた。

自由論題セッション 「環境協力評価」

座長 長谷川 弘（広島修道大学）

本セッションでは「環境国際協力評価活動に関する一考察：わが国と欧米諸国と国際機関との比較、発

表者：廣野良吉（政策研究大学院）、「北東アジアにおける経済発展段階に応じた地域環境協力の形成、発表者：青正澄（共立総合研究所）、「持続可能な開発評価の基礎的研究（2）：環境税導入国を中心とした経済・税・環境との関係の検討、発表者：泉田晃彦（国際医療福祉大学）梅内拓生（同）」及び「廃棄物分野における評価活動、発表者：和田英樹（廃棄物政策研究所）」の4件の論題が報告された。

「環境国際協力評価活動に関する一考察」では、日本、欧米諸国及び国際機関の援助事業における環境政策・評価の比較結果が報告された。環境配慮の実施時期には三者間で差があるものの内部環境評価の手順、国別援助計画実施のための環境配慮ガイドライン導入などは基本的に共通する一方、環境分野における成果評価や外部評価など日本が海外に学ぶべき点も多いことが明らかにされた。

「北東アジアにおける経済発展段階に応じた地域環境協力の形成」では、ストックホルム会議（1972年）以降のグローバルな環境協力活動を踏まえ、日中韓を中心とする北東アジア地域での持続可能な環境協力実現に向けた課題と展望が報告された。環バルト海沿岸地域を事例とし、地域環境レジームの構築あるいは認識共同体の重要性が説かれ、科学的調査・研究における本学会の役割も議論された。

「持続可能な開発評価の基礎的研究（2）」は、環境税の導入を開始している15カ国を対象に、GDP、炭酸ガス排出量、環境税収額及び総税収額の関係性についての分析結果を報告したものであった。CO₂排出量とGDPの間には従来の研究成果と同様の正の相関が認められたが、同一経済レベルの国々であってもそれらのCO₂排出量には差があり、各国の環境税の内容が排出量を大きく左右してしまう可能性が示された。

「廃棄物分野における評価活動」では、廃棄物関連事業における内外の評価活動状況と課題が浮き彫りにされた。これまでは廃棄物処理システムの事前評価が主流であったこと、行政評価の一環として国内自治体で行われてきた廃棄物関連事業評価も事務事業評価に終始し中途半端なものが多い実態が明らかにされた。模範事例として文京区の評価サイクルが示され、代替シナリオの設定、定量的評価軸、パフォーマンス評価、市民参加などの重要性が説かれ、今後のあるべき評価の方向性について活発な質疑が交わされた。

自由論題セッション 「ODA評価の課題」

座長 三好 皓一（立命館アジア太平洋大学）
 討論者 松本 哲男（名古屋大学）

本セッションでは4つの報告がなされた。

まず、三輪徳子会員（国際協力機構）から、「援助の『妥当性』の評価に関する考察」について報告がなされた。DAC評価に焦点を当て、効率的な援助実施の観点から、妥当性の評価の現状について問題を提起し、妥当性の視点の重要性と妥当性の評価のあり方について論じた。特に、妥当性については評価において重要な項目であるが十分な議論がなされていない、また、各援助機関においても認識に差異ある点が報告された。また、今後の課題として、判断基準の明確化、妥当性と効果の関係、妥当性の視点の範囲の3点が指摘された。報告に対しては、受益者に視点、時間における妥当性の変化、価値観の変化など観点から討議がなされた。源由理子会員（国際開発コンサルタント）は「開発援助事業におけるプロセス評価に関する考察 - コミュニティ開発型プロジェクトを事例として - 」について報告を行ない、ケニアの貧困層エンパワーメント・プロジェクトの終了時評価を事例として取り上げ、開発援助プロジェクトのプロセス評価を行なうことの意義を問いかけた。特に、新しい技術や知識に触発された人々が自分たちの文脈の中でそれを内在化していく「プロセス」を見ることが重要であること、また、それら进行评估することによって開発援助の戦略、すなわちアプローチや自立発展性を考察できること、さらには、援助介入の有効性を高められることの重要性について指摘がなされた。報告に対しては、オーナーシップ、成果重視などの観点から討議がなされた。永吉洋之会員（東京大学）は、「費用便益分析を用いたODA水供給プロジェク

トの効率性評価に関する研究」について報告を行なった。本報告では、ODA評価の効率性に関して、セクター毎の特徴を考慮した具体的な評価手法の必要性を指摘し、スリランカ国における水供給プロジェクトの事例を使い、有収水量増加に伴う水道事業収入増加分、施設の維持管理費削減分、貨幣価値換算した水汲み拘束時間短縮、水の購入費用削減分の4項目を指標化に基づく効率性評価の手法を提示した。報告に対しては、都市と農村における生活環境の違いが与える影響、感度分析の必要性などの観点から討議が行われた。野田邦雄会員（国際協力銀行）は、「合同評価を通じた評価体制整備への取組」について報告を行ない、JBICが実施している終了案件に対する合同評価を通してのタイ、インドネシア、インドの計画官庁、受入実施機関の評価能力の向上、評価体制整備に対する支援の試みを紹介した。合同評価の成果としては、DAC 評価5項目の理解のための技術移転、オーナーシップの強化、制度調和化を、また、合同評価のフォローアップや合同評価の管理・活用についての必要性を指摘した。報告に対しては、合同評価の利点、開発援助としての評価、合同評価のドナー指向性などの観点から討議がなされた。

自由論題セッション 「政策評価の課題」

座長 山谷 清志（同志社大学）

政策評価の課題というこのセッションの4報告は、大きく二つのテーマに区別される。一方は政策評価の現状問題に関する報告である。第1報告「公的研究機関の技術評価」で原田隆会員（産業技術総合研究所）は多額の資金を要する研究開発プロジェクトのアカウントビリティ問題について論じた。評価がその鍵になるのは当然であるが評価疲れ、一律評価の困難、評価者のレベルの3問題が自由な研究を難しくしているという。要するに、公的研究機関の技術評価は評価におけるコミュニケーションの問題である。第2報告「政策評価にみる本質的矛盾 アカウンタビリティと経営」において古川俊一会員（筑波大学）は、政策評価が求めるアカウントビリティ志向と生産性・業績志向が矛盾する点、日本の行政の制度的特徴・文化伝統がNPMを標榜するアングロサクソン国家のそれと違う点を指摘、万能薬はなく、役割を限定しつつシステム的に補完することを提言する。深い洞察による鋭い提言である。

他方のテーマは評価の取り組むべき課題である。第3報告「途上国の財政制度と財政執行の評価方法：PEFAプログラムの業績測定枠組みの考察」において上野宏会員（南山大学）は、途上国の財政制度に向けた業績測定の枠組み（PEFAプログラム）を紹介、それが先進国、とくに日本においても有効ではないかという提案をする。業績測定と予算会計制度の関連をめぐる混乱を考えると、傾聴に値する重要な提案である。最後の第4報告「日本の構造調整借款20年間のレビューからみる日本政府の政策判断」（田中弥生・東京大学）は、実務と言うよりは評価の研究として重要な検討課題である。これまでの評価は政策にしても組織運営管理にしてもミクロな部分を対象にして研究されてきたが、確かにこの報告が言うように長いスパンのマクロ政策に関する研究・評価は少なかった。一つの大きな研究領域の扉を開いた観がある。以上、この4報告はいずれも非常に刺激的であり、また今後の実務と研究にとって多くの示唆に富んだ貴重な報告であった。座長や司会者としてこれほど興奮した経験はかつてなかったことを付け加えたい。

自由論題セッション 「ODA評価手法」

座長 石田 洋子（コーエイ総合研究所）

本セッションでは「理論に基づいた開発援助の評価 発表者、山下道子（JBIC）」、「開発事業におけるインパクト評価手法改善への取り組み 発表者、小林信行（JBIC） 野田邦雄（同） 魚住麻子（同）」、「ベキ法則に基づく健康と経済のマクロ指標を用いた政府開発援助政策評価法の基礎研究（ ） 発表者、

正木朋也（統合社会科学研究会）、梅内拓生（国際医療福祉大学）、「ODA国別評価手法についての考察発表者、駒澤牧子（アースアンドヒューマンコーポレーション）、金子真知（同）」の4つの論題が報告された。

「理論に基づいた開発援助の評価」は、PRSに関連して理論の再構築を迫られる世銀の取り組みが報告された。プログラム理論再構築の手法と課題、開発プロジェクトの有効性などが主な論点となり、質疑応答では、世銀のPRS評価の試みについて具体的な内容に関心が集まった。

「開発事業におけるインパクト評価手法改善への取り組み」では、JBICのインパクト評価手法改善について、中国やインドの評価事例を通して、手法や課題が紹介された。質疑応答では、貧困削減への円借款事業の貢献、既存データに頼るのでなく計測可能な指標を取る必要性などが議論された。

「ベキ法則に基づく健康と経済のマクロ指標を用いた政府開発援助政策評価法の基礎研究（ ）」は、一人当たりGDPと乳児死亡率の関数がベキ法則に従うことに着目し、社会安定性との関連性分析の進捗報告であった。国の経済と健康の関係を評価する一助となる可能性もあり、今後の成果が期待される報告であった。

「ODA国別評価手法についての考察」は、外務省ガイドラインや先行研究に基づく政策評価手法に関する報告であった。援助額に基づいてODAの貢献度をチェックする国別評価フレームワークが提案され、外務省やJICA、JBIC評価担当部署からの参加者によって、それぞれの活動に基づくコメントや内容修正が積極的に行われた。

自由論題セッション 「保健衛生評価」

座長 梅内 拓生（国際医療福祉大学）

本セッションでは以下の4題の発表があった。

「わが国の健康づくり指標と評価 健康日本21とローカル・マニフェスト」伊関友伸（城西大学）

健康日本21の理念を行政レベルで実効性あるものにするをねらってローカル・マニフェストと連携する可能性を探る視点から、行政の取り組みを分析、評価している。このために自治体行政評価の進化モデルを考案し、行政と住民との相互協力による理念の実行レベルでの進展がローカル・マニフェストとの連携により可能になる見解をしめした。

「保健医療薬剤政策評価の基礎（2）循環器系薬剤を中心に 」高田武治、梅内拓生（国際医療福祉大学）

健康維持に必要である薬剤の適正使用の評価にはその国の政治、経済、文化などさまざまな要因を考慮しなければならない。本報告は日本と米国を比較し日本では高血圧、喫煙率が高く、米国では肥満や、カロリー摂取率が高いこと、さらに米国では疾患発生率が高いにも自分は健康であると思っている率が高いといったことなどから薬剤の適正使用への取り組みとその評価は国民、住民の健康観を視野に入れて行うことが重要であることを指摘した。

「適切な医薬品使用に関する考察（2）」梅原貞臣、梅内拓生（国際医療福祉大学）

本発表はストレス関連薬品の生産量、使用量と、自殺率や離婚率などの社会要因との関連を中心に日本と米国を比較し、医薬品適正使用を目指した行政や住民活動の評価には宗教も含め個人、社会レベルでの健康観、価値観などを視野に入れて行うことが大切であることを指摘した。

「保健医療プロジェクトの新しい学際的な評価法についてのケーススタディ」

仲佐保（国立国際医療センター）、中村安秀、草郷孝好、池田光穂、天沼直子（大阪大学）、倉沢愛子（慶応大学）

JICA保健医療プロジェクトの評価を行うために関係する多方面の専門家チーム（保健医療、医療人類学、社会学、開発経済学）による合同現地調査を実行した。その結果、途上国の多様なリアリティを複眼的にとらえる可能性があることを発表した。

自由論題セッション 「自治体による評価の課題」

座長 小野 達也（鳥取大学）

本セッションでは4つの報告が行われた。

まず、茂木勇氏（立教大学大学院）の「基礎自治体における参加型評価に関する一考察～相模原市村富公園再整備事業を事例として～」が報告された。本報告は、街区公園の再整備事業における住民参加の事例をもとに、参加型評価の可能性に言及しようとするものであった。

次に山崎肇氏（県立広島大学大学院）の「公園を介した官民連携」が報告された。本報告は、公園を介した官民連携の4事例を紹介しつつ、今後の官民連携における指定管理者制度の有用性を指摘したものである。

3番目の報告は田崎智宏氏（国立環境研究所）の「問題指摘着目型の法制度の実態評価～その類型と試行：リサイクル法制度を例に～」であった。本報告は、リサイクル法制度について問題指摘着眼型の実態評価を試みたもので、自治体における各種制度の見直しに向けた評価にも通じる論点が提出されている。

最後に、田中啓氏（静岡文化芸術大学）の「自治体評価にみる本質的課題とその解決方法に関する考察」が報告された。本報告は、自治体の行政評価が今新しい段階に差し掛かっており、これまでの考え方や取り組みの変更・転換が求められるのではないかという問題提起を行ったものであった。

なお、本セッションにおける4報告のうち3本は、セッションのタイトルである「自治体による評価の課題」そのものを直接論じたものではなかったが、報告の内容とフロアとの質疑応答は、何れも「自治体における評価」なるものに示唆を与えるものであったことを付言する。

自由論題セッション 「開発能力アセスメント」

座長 松岡 俊二（広島大学）

本セッションは、高橋与志（広島大学）・松岡俊二（広島大学）・工藤高志（広島大学）「社会的能力アセスメントに基づく援助投入評価」、登丸求己（玉川大学）「後発移行経済国タジキスタン政府の開発マネージメント・キャパシティ・アセスメント」、村上一真（広島大学）・松岡俊二（広島大学）「環境対策における社会的能力の評価」という3報告より構成された。全体として3報告は、開発政策や開発援助の前提としての当該社会のマクロ的（社会的）能力を問題とし、能力のアセスメント手法とその適用のあり方に焦点を当てたものであった。

高橋与志（広島大学）・松岡俊二（広島大学）・工藤高志（広島大学）「社会的能力アセスメントに基づく援助投入評価」は、広島大学と三菱総合研究所が行っているJICAアジア地域特定テーマ評価「経済連携事業」に基づく報告であった。報告では、マレーシア、タイ、インドネシア、フィリピンのASEAN4カ国における貿易分野のキャパシティ・デベロップメントを対象に、貿易センター事業を中心としたJICA援助を、広島大学21世紀COEプログラムが研究開発を行っている社会的能力アセスメント手法により評価を試みたものである。貿易分野の社会的能力を、アクター・ファクター分析や能力の発展ステージ分析とい

う視角から分析し、援助投入の整合性（妥当性）や援助の能力形成への貢献を評価したものであった。

登丸求己（玉川大学）「後発移行経済国タジキスタン政府の開発マネージメント・キャパシティ・アセスメント」は、報告者がJICA専門家として、タジキスタン政府の開発行政を分析し、改革の提言を行うために実施した開発管理能力のアセスメントに基づく報告であった。報告では、開発システムの整合性、開発行政機構の効率性、人的資源の生産性という3つの評価基準が提示され、開発行政の改革が提言された。

村上一真（広島大学）・松岡俊二（広島大学）「環境対策における社会的能力の評価」は、北九州市と大阪市の大気汚染対策能力の形成を対象として、政府（都市行政）部門における社会的能力の構成要素を因子分析により定量的に提示した報告であった。報告では、人材・予算などの「環境政策資源」（人的・組織的資源要素）、企業への融資などの「経済的支援」（政府・企業関係要素）、研究報告などの「科学的知見」（知識・技術要素）、立入検査などの「直接規制」（政策・対策要素）という4要素が提示された。また、人的・組織的資源要素が能力全体の5割程度を説明するという興味深い分析結果が報告された。

自由論題セッション 「教育評価」

座長 齊藤 貴浩（大学評価・学位授与機構）

本セッションでは4つの報告が行われた。

渡邊美紀（東京工業大学大学院） 牟田博光（同）による「数学的リテラシー(PISA)に影響を及ぼす国別要因の分析」では、生徒の学力到達度調査（PISA2003）における23カ国の個人データから、従来の個人、家庭、学校の各要因に加え、教育政策・施策の決定権限がどこにあるかという要因が、学力に影響を与えていることが示された。この要因がどのように学力に影響を与え、また最終的に個人、家庭、学校の各要因にブレークダウンされていくのか、それらの過程も含めたさらなる調査研究が望まれる。

HARUN, Imansyah（東京工業大学大学院） 小笠原愛美（同） 牟田博光（同）による「インドネシア中等理科教育における指導方法改善の効果」は、当該パイロット・プロジェクトの効果について論じた研究である。プロジェクトにより、実験校が統制校（比較対象となる学校）に比べて指導方法に優れ、そのことが、生徒の態度に影響を与えていることが示された。この評価を詳細に分析し、より効果的なプロジェクトとなることが期待される。

田畑直子（メルボルン大学）による「学校評価の実践：広島市の公立校4校の事例を通して」では、広島市の学校評価制度を事例とし、学校教育目標に即した学校評価によって教職員の意識改革が進み、意思統一が図られ、そして教育の質の向上に繋がるという流れが示されるとともに、留意点として、全教職員による分担、評価項目の厳選、学校の実態に沿った評価システムの確立が挙げられた。この実践事例及び帰結は他の事例とそう変わるものではなく、学校評価に関する研究は今後その普遍性と特殊性についてより検討が加えられていくものと考えられる。

高田恵子（国際医療福祉大学大学院） 梅内拓生（同）による「人間性教育と性教育評価の基礎 第2報」では、日本の青少年を取り巻く家族、社会、規範、性行動などの社会的要因について紹介があり、性感染症への感染や中絶等の指標は他国と比べて数値的には問題ではないが、その増加率等が表す社会的変化は注視されるべきであり、問題が拡大する前に、学校、家庭、行政が協力して取り組むことの必要性が報告された。次の報告では更なる進展が見られることを願う。

以上、どれも優れた報告であり、広島市の教育委員会委員長にもご参加いただくなど、フロアとの間で活発な議論が行われた。参加者の皆様には心から感謝申し上げます。

自由論題セッション 「日本評価学会最初の5年の歩み」

座長 古川 俊一（筑波大学）

本セッションでは「ロッシら『プログラム評価：システムティックアプローチ第7版』の改訂プロセスに見る米国プログラム評価研究の発展と日本における課題」（発表者、大島巖（東京大学）、「日本評価学会5年の歩みと政策評価、ODA評価」（発表者、山谷清志（同志社大学）、「自治体行政評価の動向 2001年から2005年まで」（発表者、小野達也（鳥取大学）の3つの論題が報告された。

大島報告では、1979年に体系的テキストとして出版されたロッシらの改訂経過を細かく考察し、対人サービスプログラムの評価を重視していること、1990年代以降学術引用件数が増加していること、プログラム理論、利害関係者の位置づけの変遷など翻訳者ならではの指摘がなされた。最後に日本におけるプログラム評価への示唆がなされた。

山谷報告は実際にこの5年政策評価の委員として、またODA評価の実務責任者としての経験をふまえ、歴史的な回顧をしたあと、評価学会大会報告の分析から実務上と理論・研究者のすみわけ、手法と対象にわかれた報告という分析をした。残された課題として、政策評価については、独立行政法人の評価、客観性の担保評価、あるいは政治レベルに近いODA評価などの課題が示された。また、学会の特徴から、今後学際的な特徴を伸ばしていくことが提案されるとともに、政策評価、ODA評価、独立行政法人評価の間の関係の整理についての試論が示された。

小野報告は地方自治体の評価導入の状況、評価学会における研究、発表の詳細な経緯を紹介するとともに、行政評価の課題認識と、マネジメントサイクルの実現、外部との関係構築、基礎的ツールの開発など、真価発揮の条件について述べた。

会場からは、学会が今後目指すべき方向、例えば評価文化の普及などについて、活発な意見がかわされた。

自由論題セッション 「満足度意識調査の評価活用」

座長 星野 敦子（十文字学園女子大学）

本セッションでは、コメンテーターとして齋藤貴浩先生（大学評価・学位授与機構）をお迎えして、「大学生の意識調査をベースとした日本のODAの説明責任に対する間接評価 発表者、高千穂安長」、「韓国におけるサイバー大学生の満足度に影響する要因分析及び学習環境把握 発表者 金 性希（東京工業大学） 牟田博光（同）」、「自治体公共サービス満足度ランキング向上の方法論 発表者 大嶋章嘉（市民満足学会/ワード研究所）」、「大学における授業改善をはかる要因類型化手法 発表者 星野敦子（十文字学園女子大学） 北原俊一（同） 新行内康慈（同） 安達一寿（同） 綿井雅康（同） 牟田博光（東京工業大学）の4つの論題が報告された。

「大学生の意識調査をベースとした日本のODAの説明責任に対する間接評価」は大学生に対する意識調査に基づいてODAの説明責任の実態分析を行ったもので、間接的にODAの評価をするという新たな試みについての報告であった。サンプリングや調査方法などについていくつかの提案がなされ、今後の研究の発展が望まれる。

「韓国におけるサイバー大学生の満足度に影響する要因分析及び学習環境把握」は2001年に韓国において設立されたサイバー大学の学生を対象としたアンケートに基づいて、満足度の規定要因の分析を行ったもので、多数の関連要因を取り入れた重回帰モデルを中心とした意欲的な研究である。新たな学習形態による大学評価という点で、議論も活発に行われ、更なる研究の進展が期待される。

「自治体公共サービス満足度ランキング向上の方法論」は自治体の満足度調査に基づいて、ランキング向上のためにシミュレーションを実施し、向上のための具体的な方法論を示した。千葉県浦安市のケースを取り上げ、实际的で示唆に富んだ報告であった。サービス部門の重要度をどのように決定するかという点が課題として議論された。

「大学における授業改善をはかる要因類型化手法」は大学における授業評価とコンジョイント分析を組み合わせるにより、授業評価の結果を的確に授業改善にむすびつける方法を提案した。コンジョイント分析は環境評価の分野で広く利用されていることから、環境評価への援用の可能性が議論において示唆された。

自由論題セッション 「ジェンダー評価」

座長 村松 安子（東京女子大学）
 討論者 大沢 真理（東京大学）

本セッションでは以下の4報告があった。「日本学術会議におけるジェンダー主流化への胎動」原ひろ子（城西国際大学大学院）「ジェンダー関係の平等化推進と成人学習の評価」國信潤子（愛知淑徳大学大学院）北九州市男女共同参画基本計画進捗状況評価」三隅佳子（(財)アジア女性交流研究フォーラム）山田貴宏（北九州市役所）「男女共同参画の視点から見た環境施策評価 北九州市の事例を手がかりとして」織田由紀子（(財)アジア女性交流研究フォーラム）

1995年以来国際社会で合意されている男女平等社会実現のための政策手段が「ジェンダーの主流化」である。「主流化」は政策・施策の企画段階から最後の評価までのすべての段階・局面にジェンダー視点を導入することであり、主流化それ自体が評価を含んでいる。しかし評価学の中でのジェンダー分析、特にジェンダー評価はまだ緒についたばかりである。

報告は4本とも、評価が立ち遅れている分野での革新的な報告であった。原報告は、1970年代から始まり、2000年には2010年までに女性会員10%と数値目標を設定し、2005年には「課題別委員会 学術とジェンダー」を立ち上げた日本学術会議における取組みが将来の課題と共に報告された。國信報告は、男女共同参画社会形成に不可欠な成人教育のジェンダー主流化が、最近誤った情報や情報の不正確な伝達の影響を受けてその内容が「すり替え」られている現状を受けて、有効な対抗手段として客観的成人教育評価へのDAC5段階評価の適用を提案する。三隅・山田報告は、北九州市の男女共同参画施策の評価過程を、進捗状況評価、簡易評価、詳細評価、重点評価から総合評価にいたる全プロセスに渡って報告した。全国の地方公共団体が試行錯誤中の新しい評価法確立に向けての努力である。織田報告は、民生部門、特に市民の日常生活の中からの地球温暖化対策としての「環境家計簿」データを使って、ジェンダー視点からの環境政策評価に必要なデータの検討を行った。

コメンテーターは研究報告の一層の展開のための議論だけでなく、追加情報を補足、評価指標に対する具体的提案・方法、データの単位（世帯データか個人データかなど）など多くの提案でセッションの議論をより一層興味深いものにした。

自由論題セッションX 「評価インターンによる調査報告」

座長 源 由理子（国際開発コンサルタント）

本セッションは、学会研修委員会が中心となって取り組んでいる「評価インターン制度」の中間報告を兼ねて、各インターンによる調査報告を行ったものである。

まず、広島県立教育センターについては「事業改善評価手法の提示」を目的とした調査報告が、原口裕紀子会員、山田孝子会員（以上、立教大学大学院）により発表された。評価の目的は同教育センターが事業改善評価を行うための評価手法のあり方について提示するもので、その背景には予算の適正配分と研修講座の質の向上を図るためにより適切な評価手法を導入したいという意向がある。報告では、評価手法を検討するための手段として研修事業の一部を評価する方法論と作業の途中経過が報告された。コメントーターの寺地靖仁会員（尾道市教育委員会）からは、評価方法は概ね妥当であるが今後は組織内に埋もれたデータをより活用し評価の簡素化を図ることも重要であるとの指摘があった。

次に、尾道市栗原北小学校に派遣された串本剛会員（広島大学大学院）、渡辺明美会員（早稲田大学大学院）から、「学校評価表の開発とその課題」について発表があった。同小学校が独自に開発した学校評価表とインターンが作成した評価表を比較することにより、評価表の開発・利用における課題を明らかにするもので、報告では調査結果を踏まえ重点化（簡素化）、客観性の向上、構造化を特徴とした評価表の提案がなされた。コメントーターの森田哲会員（広島県立広島中学校）からは学校評価表は学校改善のツールであるがともすると簡素化し過ぎて形骸化する恐れがある点に留意が必要であること、学校内部に評価人材が必要とのコメントがあった。フロアーからは客観性に関連し、主観的とされる観察などの手法を排除することの是非が論点として挙げられた。

（特活）メロウ倶楽部の評価では、藪並郁子会員（立教大学大学院）から事業評価の途中経過について発表があった。メロウ倶楽部は中高年層を対象としたインターネットを活用した会員貢献・社会貢献活動を行うNPOであるが、評価インターンによる事業の振り返りを通し活動を見直し、組織運営のあり方を検討することへの期待がある。コメントーターの西出順郎会員（琉球大学）、駒澤牧子会員（アースアンドヒューマンコーポレーション）から、評価の枠組み、調査のポイント、評価目的について再整理する必要性が指摘され、フロアーからは組織評価、活動評価などの言葉の使い方の再考が必要とのコメントがあった。

最後の発表となった岡山NPOセンターの事例では、佐藤浩子会員（立教大学大学院）、中川力介会員（早稲田大学大学院）から組織評価の途中報告があった。岡山NPOセンターは岡山県の指定管理者として中間支援組織の機能を持つが、組織強化のために第三者の視点から評価を行うものである。インターンが作成した組織評価フォーマットに基づき各項目ごとの評価結果（中間報告）が報告されたが、コメントーターの西出会員からは評価項目の設定、評価のプロセスにクライアントの顔が見えにくいことが指摘され、フロアーからは組織評価フォーマット作成過程及び中間支援組織の研究に関する先行研究へのコメントがあった。中間支援組織の問題点は広く調査研究されているので、今後はポジティブな面に眼を向けることにより打開策を探っていくことも必要との指摘もあった。

以上四つの報告は評価手法の開発、事業評価、組織評価と異なった評価理論と、学校評価、NPO評価の異なった分野に基づくものであるが、各コメントーター及びフロアーから適切かつ活発なコメント・指摘をいただいた。3月の最終報告に向けて、これらのコメントを反映したより質の高い評価調査の実施を期待したい。

企画委員会

1. 研究受託事業

平成17年度は国際協力機構より「事業評価の二次評価」について受託した。国際協力機構はすべての技術プロジェクト案件について終了時評価を行っているが、年間約50件とその数が多いところから、内部評価として実施している。しかし、評価内容の透明性、客観性を担保するために、作成された評価報告書を外部有識者事業評価委員会で確認する事になっている。外部有識者事業評価委員会の下に作業部会を設け、計量的な分析を加味して、公正なレビュー - を行っているが、この部会の仕事を学会が請け負ったものである。

この成果の一部は平成17年12月開催の第6回全国大会で「外部有識者によるODA評価の二次評価」(共通論題セッションV 評価の評価 評価の質保証のための仕組みと手法)として一部発表された。また、この結果は平成17年度末に公表された国際協力機構「事業評価年次報告書2005」の中に集録されている。

このような受託事業は学会としての社会的貢献を具体的な形で示すと共に、会員に評価作業を経験してもらう場を提供することにもなる。さらに学会に対しては財政的な貢献もできるなど利点大きい。企画委員会としてはこのような受託事業を積極的に拡大していきたいと考えている。

2. 顕彰事業

学会賞には論文賞、奨励賞、功績賞の3種がある。

2005年度は学会賞受賞者選考細則第6条により、2005年8月30日付けで学会ホ - ムペ - ジ上に学会賞候補者の公募を行い、11月10日まで自薦、他薦を受け付けた。11月22日に学会賞審査委員会を開催した。論文賞、奨励賞、功績賞のそれぞれについて、申し合わせに基づき、功績、論文、発表原稿等公表された資料、アンケート結果、及び、日頃の学会活動などを総合的に検討し、さらに、受賞候補者の意思確認のためメルにて議論を継続し、12月5日、学会賞受賞候補者選考について審査員全員一致で下記の結論に達した。

学会賞受賞者選考細則第7条では、論文賞は、選考当該年度の前年度に発行された本学会誌に掲載された原著論文等から推薦されるものとする、と定められている。日本評価研究第5巻第1号及び第2号の原著論文の中から、編集委員会の推薦に基づき、論文賞受賞候補者として、大久保智哉会員を選考した。

奨励賞は、選考当該年度の前年度に開催された全国大会での発表等から推薦されるものとする、と定められている。第5回全国大会、及び春季第2回全国大会の場で発表された原稿等公表された資料、及び、それぞれの大会でのアンケート結果などを総合的に検討し、受賞候補者として田中弥生会員を選考した。

功績賞は、今回は該当者なしとした。

これらの選考結果は、企画委員長を通じて理事会に諮られ、原案通り承認された。また、この2人の受賞者には、12月10日の総会の場で、学会から賞状及び副賞が授与された。

企画委員長 牟田 博光

研修委員会

活動方針：

研修委員会では、「評価の普及および評価に関する人材育成に寄与すること」を目的として、各種セミナーの開催、研修プログラムの開発、教材の作成、講師の派遣、その他人材育成に関わる活動の実施支援を行うことを基本方針とし、2005年度には以下のような活動を展開した。

活動状況：

1. 「評価インターン出前サービス」の企画・実施

「評価インターン出前サービス」は、学会の学生会員に対して「実際に評価を経験する」機会を提供すべく、「評価を外部に委託したい」と考えている地方自治体やNGO・NPOで実際に評価を体験させていただく事業である。今年度は、以下の4団体が受け入れ団体となり、計7人の学生会員をインターンとして派遣した。その途中報告は2005年12月の全国大会（広島大学）の場で報告され、2006年3月には、最終報告書がとりまとめられた。

- 1) 広島県立教育センター 「広島県立教育センターの事業改善評価」
山田 孝子（立教大学21世紀社会デザイン研究科）
原口裕紀子（立教大学21世紀社会デザイン研究科）
- 2) 広島県尾道市立栗原北小学校 「学校評価表の開発とその課題」
串本剛（広島大学教育学研究科）
渡辺明美（早稲田大学アジア太平洋研究科）
- 3) NPO法人メロウ倶楽部 「メロウ倶楽部の事業評価」
藪並郁子（立教大学異文化コミュニケーション研究科）
- 4) NPO法人岡山NPO活動センター 「岡山NPOセンターの組織評価」
中川力介（早稲田大学公共経営研究科）
佐藤浩子（立教大学21世紀社会デザイン研究科）

研修委員会共同委員長 源 由理子
西出 順郎

広報委員会

1. 活動方針

広報委員会は、その目的である、「評価の意義を広く普及し、学会活動の強化を支援するために、日本評価学会および学会の行う種々の事業を広報し、学会の情報活動に寄与する」ことを、したがって、評価学会の活動のみならず広く評価に関わる情報も会員の皆さんと共有すべくあらゆる機会を捉えて実現を目指します。

2. 活動状況

(1)「学会活動の強化と学会の健全な運営を目指して会員獲得に努めます」

会員獲得のための学会説明などのツールについては、ホームページを第一に考えていますが、国、自治体、企業などの説明ツールとしては、マス・メディアを使ったパブリシティを利用する方向で引き続き交渉します。

(2)「会員に対する広報活動の充実」

外務省、独立行政法人国際協力機構（JICA）の評価報告をタイムリーに評価学会会員に広報できるよう、外務省、国際協力機構の評価資料を公表後、内容について転送する仕組みを構築しました。

(3)「広報の手段は主として学会ホームページを活用します。従って、学会からのお知らせはホームページに随時掲載します」。

より読みやすく、タイムリーに情報をお伝えすべくホームページリニューアルの外注を計画し、業者を選定しましたが、具体的な内容について引き続き交渉中です。

(4)「ホームページの有効性を評価します」

2006年には、ホームページの有効性評価を行うために、ホームページへのアクセス件数、アクセス件数を会員数で除した数値の検討などを実施します。

3. 今後の方針

広報委員会は、第3回委員会を2005年12月28日に開催しました。引き続き年2回の広報委員会の開催とともに、進捗状況管理、新企画の検討などを行い広報面での有効性、効率性を高めて行きます。

広報委員長 高千穂 安長

国際交流委員会

1. 地球環境に関する合同ワークショップ

地球環境ファシリティ（GEF）は途上国における地球環境プロジェクトに世界銀行、UNDP、UNEPなどを通じて無償で資金供与を行う国際機関であり、日本はその主な出資者である。わが国におけるGEFならびに地球環境問題に対する認識を深めるため、昨年11月23日、当委員会主催・GEFおよび国連大学共催による合同ワークショップが東京青山の国連大学エリザベス・ローズ会議場で開催された。

当日、財務省（宮原国際局開発企画官）、環境省（田中地球環境局環境協力室長）の挨拶に続き、Zakri 国連大学高等研究所長、廣野日本評価学会副会長、Good地球環境ファシリティ事務局長の基調演説（代読）に続いて、今年度のGEFの活動の評価を担当したICF Consulting社 のWagner上級副社長から評価報告書である"Third Overall Performance Study of the Global Environment Facility"（OPS3）の説明、柳上智大学大学院教授の「日本の対中環境協力プログラム」の分析評価、Dr. Tengberg UNEP Programme Officerの「GEFを通じたUNEP/国連大学のプログラム協力」の説明があった。午後には、再びWagner氏からOPS3 の評価手法についての説明、国連大学および同高等研究所所員から国連大学の環境プログラムについての説明がなされ、最後にGEFの活動の評価と今後の課題や展望についてのセッションが持たれた。各セッションの議長は、廣野氏、Zakari氏、Van den Berg GEF評価室長、安井国連大学副学長が交互に務めた。

前号にて報告の通り、当委員会は昨年8月以降、分科会にわかれてOPS 3などの分析を進めてきた。11月には全体会議を開き、GEFの活動についての当委員会の評価と提言の内容を決めた。本ワークショップへの当委員会からの出席者は、廣野氏の他、梅内氏（国際医療福祉大学）、湊氏（FASID）、山下氏（JBIC）、田中氏（モンゴリア・リンクス）、藤本氏（あかう）、正木氏（統計社会科学研究会）、山口氏（コムシスネット）、山田氏（立教大学大学院生）であった。各出席者は質問・意見表明に活発に参加した。最後のセッションで、当委員会の提言として、藤本氏が次の趣旨の発言を行った。

GEFの現行の方針では、資金供与は急成長中の大国に集中し、中小途上国には恩恵が及ばない。しかし、中小途上国の環境問題も地球環境にとって無視できない。

中小途上国の政府には環境問題に関心を払う余裕はない。一方、これらの国の民間団体・私企業の中には、環境改善と経済成長を同時にもちたらず事業に強い関心と識見を持っている例がみられる（例 2004年ノーベル平和賞受賞者ワンガリ・マータイ女史のGreen Belt Movement）。しかしこれらの民間組織は国内において無視される傾向がある。

これらの民間組織の希求と識見を環境改善と経済成長に結びつけるためには、GEFが国際金融公社などの協力を得て、彼らを認知し、資金的技術的援助に加えて現地政府説得の支援を行うことが適当である。

GEFは既にSmall Grants Programを通じてNGO/NPO民間組織へも資金供与を行っているが、これには限度がある。日本評価学会は、GEFが民間組織への資金供与のより大きな支援プログラムを設定すること、そのための資金を民間からも調達する手段を開発すること、を提案する。

日本が貢献する国際機関の評価に本学会が参加することは、大いに意味があると考える。

2. 第5回ODA評価ワークショップへの参加

2006年1月26、27日にODA評価ワークショップが外務省により開催された。アジア諸国、国際機関、援助機関の参加のもと活発な議論が行われたが、日本評価学会の会員も数名が参加した。詳しくは外務省ホームページ参照。

<http://www.mofa.go.jp/mofa/gaiko/oda/index/kaikaku/hyoka.html>

「評価に関するセミナー・ワークショップ」

「第5回ODA評価東京ワークショップの開催」を参照

国際交流委員長 廣野 良吉

日本評価学会春季第3回全国大会のご案内

春季第3回全国大会

実行委員会委員長 高橋 一生

プログラム委員会委員長 牟田 博光

来たる2006年6月24日(土)、日本評価学会春季第3回全国大会が国際基督教大学(東京都三鷹市)にて開催されます。

参加申込方法および会場の詳細につきましては、決まり次第追ってご連絡いたします。またプログラムにつきましては、更新の都度、学会メーリングリストおよびホームページでご案内いたします。

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

記

1. 日時： 2006年6月24日(土)
2. 場所： 国際基督教大学(東京都三鷹市)
住所： 東京都三鷹市大沢3-10-2

以上

特定非営利活動法人日本評価学会定款

第1章 総則

(名称)

第1条 本法人は、特定非営利活動法人日本評価学会と称し、英語名は、The Japan Evaluation Society (略称JES)とする。

(事務所)

第2条 本法人は、主たる事務所を東京都江東区富岡二丁目9番11号に置く。

(目的)

第3条 本法人は、広く国民一般に対する透明性確保、説明責任遂行のため、公共的な活動やNPO等の社会的な活動に関する評価の重要性が強く認識されていることに鑑み、評価に関する研究及び応用を促進し、評価に関する関連機関等との情報交換を図るとともに、この分野の学問の進歩発展及び評価に携わる人材の育成等に貢献することにより、広く一般市民に対し評価活動の向上と評価の普及に寄与することを目的とする。

(特定非営利活動の種類)

第4条 本法人は、前条の目的を達成するため、次の種類の特定非営利活動を行う。

- (1) 社会教育の推進を図る活動
- (2) 学術、文化、芸術又はスポーツの振興を図る活動
- (3) 国際協力の活動
- (4) 男女共同参画社会の形成の促進を図る活動
- (5) 経済活動の活性化を図る活動

(事業の種類)

第5条 本法人は、第3条の目的を達成するため、特定非営利活動に係る事業として次の事業を行う。

- (1) 評価に関する調査・研究事業
広く一般市民の参加を得て客観性を保持し得る評価手法、評価基準の開発を行う。
- (2) 評価に関する研修事業
評価に関する学術研究会、講演会、国際シンポジウム、研修などを開催し、評価人材の育成に努める。
- (3) 評価に関する交流事業
評価研究者、評価実務者、評価利用者等評価に関連する内外の人々との交流、情報交換の場として国際交流や全国大会を実施する。
- (4) 評価に関する広報・普及事業
評価に関する情報を広く知らしむべく、ホームページの作成、学会誌その他刊行物の発行等、普及啓蒙活動を行う。
- (5) 評価実施事業
公平、中立の観点に立ち、法人として第三者評価を実施する。
- (6) 評価に関する顕彰事業
評価に関する研究および評価活動の普及に貢献した人を広く顕彰する。
- (7) その他、本法人の目的を達成するために必要な事業

第2章 会員

(種別)

- 第6条 本法人の会員の種別は、次の4種とし、正会員をもって特定非営利活動促進法（以下「法」という。）上の社員とする。
- (1) 正会員 本法人の目的に賛同して入会した個人。
 - (2) 学生会員 本法人の目的に賛同して入会した、原則として大学以上の学生で、学生会員を希望する者。尚学生会員は、卒業時に本人が希望する場合は会長に申し込みをおこなったうえで、正会員となることができる。
 - (3) 賛助会員 本法人の目的に賛同し、その事業を後援する団体。
 - (4) 名誉会員 本法人に功労のあった者及び広く評価分野に関連ある分野における学識経験者で理事会の推薦に基づき総会の承認を経た個人。

(入会)

- 第7条 会員の入会について、特に条件は定めない。
- 2 正会員、学生会員、賛助会員として入会しようとする者は、会長が別に定める入会申込書により会長に申し込むものとする。また、名誉会員として推薦された者は、入会の手続きを要せず、本人の承諾をもって会員となるものとする。
- 3 理事会は、前項の申し込みがあったとき、正当な理由がない限り、入会を認めなければならない。
- 4 会長は、第2項のものを入会を認めないときは、速やかに理由を付した書面をもって本人にその旨を通知しなければならない。

(会費)

- 第8条 正会員、学生会員、賛助会員は、総会において別に定める会費を納入しなければならない。

(会員の資格の喪失)

- 第9条 会員が次の各号の一に該当する場合には、その資格を喪失する。
- (1) 退会届の提出をしたとき。
 - (2) 本人が死亡し、若しくは失踪宣告を受け、又は会員である団体が消滅したとき。
 - (3) 継続して2年以上会費を滞納したとき。
 - (4) 除名されたとき。

(退会)

- 第10条 会員は、会長が別に定める退会届を会長に提出して、任意に退会することができる。

(除名)

- 第11条 会員が次の各号の一に該当する場合には、総会の議決により、これを除名することができる。
- (1) この定款に違反したとき。
 - (2) 本法人の名誉を毀損し又は設立の趣旨に反する行為を行なったことにより、総会において、除名すべきものと認められた場合
- 2 前項の規定により会員を除名しようとする場合は、議決の前に当該会員に弁明の機会を与えなければならない。

(抛出品の不返還)

第 1 2 条 本法人は、会員がその資格を喪失しても、既に納入した会費その他の抛出品は返還しない。

第 3 章 役員

(種別及び定数)

第 1 3 条 本法人には、次の役員を置く。

(1) 理事 20名以上25名以内

(2) 監事 1名以上2名以内

2 理事のうち1人を会長とし、若干名を副会長とする。

(選任等)

第 1 4 条 理事及び監事は、総会において、正会員の中から選任する。

2 会長及び副会長は、理事会において、理事の互選により定める。

3 役員のうちには、それぞれの役員について、その配偶者若しくは三親等以内の親族が一人を超えて含まれ、又は当該役員並びにその配偶者及び三親等以内の親族が役員の総数の3分の1を超えて含まれることになってはならない。

4 法第20条各号のいずれかに該当する者は、本法人の役員になることができない。

5 理事及び監事は、相互に兼ねることができず、本法人の職員を兼ねてはならない。

(職務)

第 1 5 条 会長は、本法人を代表し、会務を統括する。

2 副会長は、会長を補佐し、会長に事故があるとき又は欠けたときは、会長があらかじめ指名した順序で、その職務を代行する。

3 理事は理事会を構成し、この定款の定め及び総会または理事会の議決に基づき、会務の執行を決定する。

4 監事は、次に掲げる職務を行う。

(1) 理事の会務執行の状況を監査すること。

(2) 本法人の財産の状況を監査すること。

(3) 前2号の規定による監査の結果、本法人の業務又は財産に関し不正の行為又は法令若しくは定款に違反する重大な事実があることを発見した場合には、これを総会又は所轄庁に報告すること。

(4) 前号の報告をするために必要がある場合には、総会を招集すること。

(5) 理事の会務執行の状況又はこの法人の財産の状況について、理事に意見を述べること。

(任期等)

第 1 6 条 役員任期は、2年とする。ただし再任を妨げない。

2 補欠又は増員により就任した役員任期は、前項本文の規定にかかわらず、前任者又は現任者の残存期間とする。

3 役員は、辞任または任期満了後においても、後任者が就任するまでは、その職務を行なわなければならない。

4 前項の規定にかかわらず、後任の役員が選任されていない場合には、任期の末日後最初の総会が終結するまでその任期を伸長する。

(欠員補充)

第17条 理事又は監事のうち、その定数の3分の1を超える者が欠けたときは遅滞なくこれを補充しなければならない。

(解任)

第18条 役員が次の各号の一に該当する場合には、理事は理事会の議決により、監事は総会の議決により、これを解任することができる。

(1) 心身の故障のため、職務の遂行の堪えないと認められるとき。

(2) 職務上の義務違反その他役員としてふさわしくない行為があったとき。

2 前項の規定により役員を解任しようとする場合は、議決の前に当該役員に弁明の機会を与えなければならない。

(費用弁償)

第19条 役員には、その職務を執行するために要した費用を弁償することができる。

2 前1項に関し必要な事項は、総会の議決を経て、会長が別に定める。

(顧問)

第20条 本法人に顧問若干名を置くことができる。

2 顧問は、会長が理事会の同意を得て委嘱する。

3 顧問は、本法人の運営に関し、会長の諮問に答え、又は意見を述べることができる。

4 顧問の任期については、第16条第1項の規定を準用する。

第4章 会議

(種別)

第21条 本法人の会議は、総会および理事会の2種とする。

2 総会は、通常総会及び臨時総会とする。

(総会の構成)

第22条 総会は、正会員をもって構成する。

(総会の権能)

第23条 総会は、以下の事項について議決する。

(1) 定款の変更

(2) 解散及び合併

(3) 事業計画及び収支予算

(4) 会員の除名

(5) 事業報告及び収支決算

(6) 役員の選任又は監事の解任、役員の職務

(7) 会費

(8) 解散における残余財産の帰属先

(9) その他運営に関する重要事項

(総会の開催)

第 2 4 条 通常総会は、毎年1回開催する。

2 臨時総会は、次に掲げる場合に開催する。

- (1) 理事会が必要と認め、招集の請求をしたとき。
- (2) 正会員総数の5分の1以上から会議の目的を記載した書面により招集の請求があったとき。
- (3) 監事が第15条第4項第4号の規定に基づいて招集するとき。

(総会の招集)

第 2 5 条 総会は、前条第2項第3号の場合を除いて、会長が招集する。

2 会長は、前条第2項第1号及び第2号の規定による請求があったときは、その日から30日以内に臨時総会を招集しなければならない。

3 総会を招集する場合には会議の日時、場所、目的及び審議事項を記載した書面により、開催の日の少なくとも5日前までに通知しなければならない。

(総会の議長)

第 2 6 条 総会の議長は、会長をもってこれにあてる。

(総会の定足数)

第 2 7 条 総会は、正会員総数の2分の1以上の出席が無ければ開催することができない。

(総会の議決)

第 2 8 条 総会における議決事項は、第25条第3項の規定によってあらかじめ通知した事項とする。

2 総会の議事は、この定款に規定するもののほか、出席した正会員の過半数をもって決し、可否同数のときは、議長の決するところによる。

(総会での表決権等)

第 2 9 条 各正会員の表決権は平等なものとする。

2 やむを得ない理由により総会に出席できない正会員は、あらかじめ通知された事項について、書面をもって表決し、又は他の正会員を代理人として表決を委任することができる。

3 前項の代理人は、代表権を証する書面を総会毎に議長に提出しなければならない。

4 第1項の場合において、正会員は、表決内容等について、総会の議長に一任することができる。

5 第1項及び前項の規定により、表決権を行使する場合は、当該正会員は総会に出席したものとみなす。

6 総会の議決について、特別の利害関係を有する正会員は、その議事の議決に加わることはできない。

(総会の議事録)

第 3 0 条 総会の議事については、次の事項を記載した議事録を作成しなければならない。

- (1) 日時及び場所
- (2) 正会員総数及び出席者数 (書面表決者又は表決委任者がある場合にあっては、その数を付記すること。)
- (3) 審議事項
- (4) 議事の経過の概要及び議決の結果

(5) 議事録署名人の選任に関する事項

- 2 議事録には、議長及び総会において選任された議事録署名人2名が、記名押印又は署名しなければならない。

(理事会の構成)

- 第31条 理事会は、理事をもって構成する。ただし、監事、顧問、幹事ならびに第53条第2項に定める事務局長は、理事会に出席し意見を述べることができる。

(理事会の権能)

- 第32条 理事会は、この定款に別に定める事項のほか、次の事項を議決する。

- (1) 総会に付議すべき事項
- (2) 総会の議決した事項の執行に関する事項
- (3) その他総会の議決を要しない会務の執行に関する事項

(理事会の開催)

- 第33条 定例理事会は、毎年3回開催する。

- 2 前項にかかわらず次に掲げる場合には、臨時理事会を開催することができる。
- (1) 会長が必要と認めるとき。
 - (2) 理事総数の2分の1以上から理事会の目的である事項を記載した書面により招集の請求があったとき。
- 3 理事会は、理事現在数の3分の1以上の出席がなければ、開会することができない。

(理事会の招集)

- 第34条 理事会は、会長が招集する。

- 2 会長は、前条第2項第2号の場合にはその日から20日以内に理事会を招集しなければならない。
- 3 会長は、緊急に理事会を招集する必要がある場合において、やむを得ない事情によりこれを開催できないときには、理事の承諾を得て、書面により議決を得ることができる。この場合、理事会は開催されたものとみなす。
- 4 理事会を招集するときは、会議の日時、場所、目的及び審議事項を記載した書面により、開催の日の少なくとも5日前までに通知しなければならない。

(理事会の議長)

- 第35条 理事会の議長は、会長がこれにあたる。

(理事会の議決)

- 第36条 理事会における議決事項は第34条第4項の規定によってあらかじめ通知した事項とする。

- 2 理事会の議決は、出席理事の過半数の同意でこれを決し、可否同数のときは、議長の決すところによる。

(理事会の表決権等)

- 第37条 各理事の表決権は、平等なものとする。

- 2 やむを得ない理由のため、理事会に出席できない理事は、あらかじめ通知された事項について書面又は代理人をもって表決することができる。

- 3 前項の場合において、理事は、表決内容等について理事会の議長に一任することができる。
- 4 第1項及び前項の規定により、表決権を行使する場合は、当該理事は理事会に出席したものとみなす。
- 5 理事会の議決について、特別の利害関係を有する理事は、その議事の議決に加わることができない。

(理事会の議事録)

第38条 理事会の議事については、次の事項を記載した議事録を作成しなければならない。

- (1) 日時及び場所
 - (2) 理事総数、出席者数及び出席者氏名(書面表決者にあつては、その旨を付記すること。)
 - (3) 審議事項
 - (4) 議事の経過の概要及び議決の結果
 - (5) 議事録署名人の選任に関する事項
- 2 議事録には、議長及び総会において選任された議事録署名人2名が、記名押印又は署名しなければならない。

第5章 資産

(資産の構成)

第39条 本法人の資産は、次の各号に掲げるものをもって構成する。

- (1) 設立当初の財産目録に記載された財産
- (2) 会費
- (3) 寄付金品
- (4) 財産から生ずる収入
- (5) 事業に伴う収入
- (6) その他の収入

(資産の管理)

第40条 本法人の資産は、会長が管理し、その方法は理事会の議決による。

- 2 この法人の資産は特定非営利活動に係る事業に関する資産とする。

第6章 会計

(会計の原則)

第41条 本法人の会計は、法第27条各号に掲げる原則に従って行わなければならない。

(会計区分)

第42条 この法人の会計は、特定非営利活動に係る事業会計とする。

(事業年度)

第43条 本法人の事業年度は、毎年10月1日に始まり翌年9月30日に終わる。

(事業計画及び予算)

- 第44条 本法人の事業計画書及び収支予算書は、会長が作成し、理事会の議決を経た後、毎事業年度の開始前に総会の議決を得なければならない。ただし、やむを得ない事情により、当該事業年度開始前に総会を開催できない場合にあっては、理事会の議決によることを妨げない。この場合、当該事業年度の開始の日から90日以内に総会の議決を得るものとする。
- 2 会長は、前項の事業計画書及び収支予算書を変更しようとするときは、理事会の議決を得なければならない。

(暫定予算)

- 第45条 第43条の規定にかかわらず、やむを得ない理由により予算が成立しないときは、会長は、理事会の議決を経て、予算成立の日まで前事業年度の予算に準じ収入支出することができる。
- 2 前項の収入支出は、新たに成立した予算の収入支出とみなす。

(予備費)

- 第46条 予算超過又は予算外の支出に充てるため、予算中に予備費を設けることができる。
- 2 予備費を使用するときは、理事会の議決を経なければならない。

(事業報告及び決算)

- 第47条 本法人の事業報告書、財産目録、貸借対照表及び収支決算書等決算に関する書類は、会長が毎事業年度終了後遅滞なくこれを作成し、監事の監査を受け、理事会の議決を経た後、当該事業年度終了後90日以内に総会の決議を得なければならない。
- 2 決算上剰余金を生じたときは、次事業年度に繰り越すものとする。

(臨機の措置)

- 第48条 予算をもって定めるもののほか、借入金の借入れその他新たな義務の負担をし、又は権利の放棄をしようとするときは、理事会の議決を経なければならない。

第7章 定款の変更、解散及び合併

(定款の変更)

- 第49条 本法人が定款を変更しようとするときは、総会に出席した正会員の4分の3以上の多数による議決を経、かつ、法第25条第3項に規定する軽微な事項を除いて所轄庁の認証を得なければならない。

(解散)

- 第50条 この法人は、次に掲げる事由により解散する。
- (1) 総会の決議
 - (2) 目的とする特定非営利活動に係る事業の成功の不能
 - (3) 正会員の欠亡
 - (4) 合併
 - (5) 破産
 - (6) 所轄庁による設立の認証の取消し
- 2 前項第1号の事由により本法人が解散するときは、正会員総数の4分の3以上の承諾を得なければ

ばならない。

3 第1項第2号の事由により解散するときは、所轄庁の認定を得なければならない。

(残余財産の帰属)

第51条 本法人が解散（合併又は破産による解散を除く。）したときに残存する財産は、法第11条第3項に掲げる者のうち、総会において議決した者に譲渡するものとする。

(合併)

第52条 本法人が合併しようとするときは、総会において正会員総数の4分の3以上の議決を経、かつ所轄庁の認証を得なければならない。

第8章 公告

(公告の方法)

第53条 本法人の公告は、この法人の掲示場に提示するとともに、官報に掲載して行う。

第9章 事務局

(事務局の設置)

第54条 本法人に、本法人の事務を処理するため、事務局を設置する。

2 事務局には、理事会の同意を得て会長が委嘱する事務局長を置く。

3 事務局組織、運営、職員及びその他必要な事項は理事会の議決を経て別に定める。

第10章 雑則

(細則)

第55条 この定款の施行について必要な細則は、理事会の議決を経て会長がこれを定める。

(委員会及び分科会)

第56条 本法人は、事業の円滑な遂行を図るため、委員会及び分科会を設けることができる。

2 委員会及び分科会は、その目的とする事項について、調査及び研究し、または審議する。

3 委員会及び分科会の組織、構成及び運営その他必要な事項は、理事会の議決を経て、別に定める。

(支部)

第57条 本法人は、事業の円滑な遂行を図るため、総会の議決に基づいて支部を置くことができる。

2 支部には、理事会の同意を得て会長が委嘱する支部長を置く。

3 支部の位置、組織、運営その他必要な事項は、理事会の議決を経て、別に定める。

附則

1 この定款は、本法人の成立の日から施行する。

2 本法人の設立当初の役員は、別表の通りとする。

- 3 本法人の設立当初の役員の任期は、第16条第1項の規定にかかわらず、平成16年度の事業報告書及び収支決算の決議をする総会の日までとする。
- 4 本法人の設立当初の事業年度は、第43条の規定にかかわらず、本法人成立の日から平成16年9月30日までとする。
- 5 本法人の設立初年度の事業計画及び収支予算は、第44条の規定にかかわらず、設立総会の定めるところによる。
- 6 本法人の設立当初の会費は、第8条の規定に関わらず、次に掲げる額とする。

正会員	15,000円
学生会員	5,000円
賛助会員	1口100,000円（1口以上）

【別表】

設立当初の役員

役 職 名	氏 名
会 長	宮内 義彦
副 会 長	廣野 良吉
副 会 長	牟田 博光
理 事	竹内 正興
理 事	荒木 光弥
理 事	梅内 拓生
理 事	喜多 悦子
理 事	木村 孟
理 事	熊岡 路矢
理 事	坂元 浩一
理 事	白須 孝
理 事	杉下 恒夫
理 事	鈴木 絲子
理 事	高千穂安長
理 事	高橋 一生
理 事	長尾 眞文
理 事	古川 俊一
理 事	三好 皓一
理 事	森 茂子
理 事	弓削 昭子
監 事	高瀬 国雄
監 事	松下 弘

日本評価学会誌刊行規定

2005.2.15改訂

2002.9.18改訂

2001.9.9改訂

(目的および名称)

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

(編集委員会)

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

(編集方針)

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

集委員会からの学会会員への連絡等が含まれる。

13. 論文等は2名の査読者により査読することとし、その人選は編集委員会が行う。「研究論文」については、査読結果と編集委員会が査読者とは別に指名する担当編集委員1名の参考意見をもとに、編集委員会が掲載に関する決定を行う。「総説」、「研究ノート」、「実践・調査報告」および「その他」の論文については、査読結果にもとづき編集委員会が掲載に関する決定を行う。
14. 編集委員が「評価研究」に投稿した場合には、当該委員はその投稿に係わる常任編集委員会あるいは編集委員会の議事に一切参加しないものとする。
15. 上記いずれのカテゴリーの投稿についても、常任編集委員会による掲載の判断は可・不可の二者択一で行うこととする。但し、場合によっては編集委員会の判断で、小規模の修正による掲載も認める。「研究論文」としての掲載が適当でないとは判断された場合でも、投稿者が希望すれば、常任編集委員会は「研究ノート」あるいは「実践・調査報告」としての掲載を決定できる。

(投稿要領の作成公表)

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

(配布先)

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

(抜刷の配付)

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

(インターネット上の公開)

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

(著作権)

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

(事務局)

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

(以上)

『日本評価研究』投稿規定

2003.4.18改訂

2002.3.25改訂

2001.9.9改訂

1. 『日本評価研究』(The Japanese Journal of Evaluation Studies)は、評価に関する論文、論考、調査報告等を掲載する。
2. 『日本評価研究』は、会員間の研究成果交流の場を提供し、内外における評価研究の一層の発展に資することを主目的として発行されており、原則として会員による寄稿を掲載する。なお、依頼原稿を除き、ファーストオーサーは学会員でなければならない。
3. 投稿された原稿は、編集委員会の責任において審査を行ない、採否を決定する。審査にあたっては、1原稿毎に2名の査読者を選定し、査読結果を参考にする。(査読者には、投稿者名を伏せて査読を依頼する。)
4. 原稿料は支払わない。
5. 『日本評価研究』に掲載された論文等は、その全文をインターネット上の本学会のホームページに掲載する。
6. 投稿にあたっては、投稿原稿が、研究論文、総説、研究ノート、実践・調査報告、その他のうち、どのカテゴリーに入るかを明記する。ただし、カテゴリーについての最終判断は、編集委員会で行なう。「研究論文」は評価の理論構築あるいは慣行の理解について重要な学問的貢献となると認められる論文、「総説」は、評価の理論あるいは慣行について概観する論文、「研究ノート」は「研究論文」作成過程での理論的あるいは経験的な研究の中間的成果物に相当する論考、「実践・調査報告」は評価事業の実践あるいは評価にかかわる調査の報告、「その他」は編集委員会が独自に企画する特集に掲載する依頼原稿等である。
7. 投稿方法
 - (1) 使用言語は日本語又は英語とする。
 - (2) 著者校正は原則として第一校までとする。
 - (3) 英文原稿については、ネイティブスピーカーによる英文チェックを済ませ、完全な英文にして投稿すること。
 - (4) ハードコピー4部(A4版)を提出する。その際、連絡先(住所、Tel、Fax、Email)と原稿の種類を明記すること。掲載可と判断された原稿については、必要なリライトを経た後に、最終原稿のハードコピー2部とDOS/Vフォーマットのフロッピーを用いたTEXTファイルを提出する。その際、オリジナル図表を添付すること。
 - (5) 刷り上がりは最大14ページとする。これを超える場合は、その経費は著者負担とする。
 - (6) 日本語原稿の最大文字数は以下のとおり。研究論文20,000字、総説15,000字、研究ノート15,000字、実践・調査報告20,000字、その他適宜。それぞれ和文要旨を400字程度、英文要旨を150words程度、及び和文・英文でキーワード(5つ以内)を別に添付する。印刷は1ページ、

20字×43行×2段(1,720字)とする。20,000字の原稿の場合、単純計算では英文要旨1ページを加えて合計13ページとなるが、図表の量によっては、それ以上のページ数となり得るので、注意すること。

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

8. 送付先

〒135-0047 東京都江東区富岡2-9-11 京福ビル
財団法人 国際開発センター内 日本評価学会事務局
TEL: 03-3630-8031
FAX: 03-3630-8095
E-mail: jes_secr@idcj.or.jp

『日本評価研究』執筆要領

2002.9.18改訂
2002.3.25改訂

1. 本文、図表、注記、参考文献等

(1) 論文等の記載は次の順序とする。

日本語原稿の場合

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

第2ページ以下：本文、謝辞あるいは付記、注記、参考文献

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

英文原稿の場合

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

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

(2) 本文の区分は以下のようにする。

例1（日本語）

- 1 .
- (1)
- (2)
- (3)

例2（英文）

- 1.
- 1.1
- 1.1.1
- 1.1.2

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

例1：日本語原稿の場合

図1 州における生徒数の推移



- (注)
- (出所)

表 1 州における事故件数



(注)
(出所)

例 2 : 英文原稿の場合



Figure 1 Number of Students in the State of

Note:
Source:

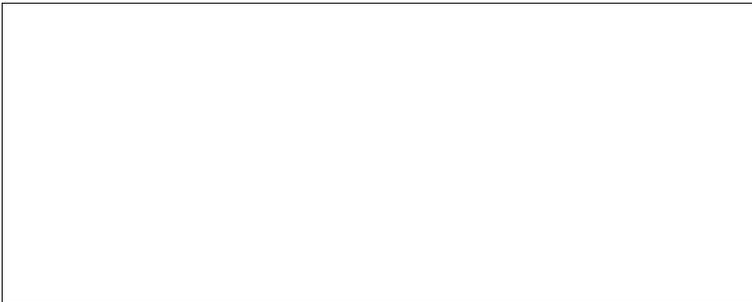


Table 1 Number of Accidents in the State of

Note:
Source:

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

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

(6) 注記、参考文献は論文末に一括掲載する。

注記

1

2

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

日本語単行本：著者（発行年）『書名』、発行所

（例）日本太郎（1999）『これからの評価手法』、日本出版社

日本語雑誌論文：著者（発行年）「題名」、『雑誌名』、巻（号）：頁 - 頁

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

日本語単行本中の論文：著者（発行年）「題名」、編者『書名』、発行所、頁 - 頁

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

複数の著者による日本語文献：著者・著者（発行年）『書名』、発行所

（例）日本太郎・日本花子（2002）『政策評価』、日本出版社

英文単行本：著者（発行年）。書名。発行地：発行所。

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

英語雑誌論文：著者（発行年）。題名。雑誌名、巻(号)、頁 - 頁。

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

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

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

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

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

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

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

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

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

日本評価研究査読要領

日本評価学会 『日本評価研究』編集委員会
2005年9月10日決定

1. 本査読要領の趣旨

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

2. 査読の目的と投稿者の責任

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

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

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

3. 査読の視点

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

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

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

4. 投稿に当たっての留意点

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

論文等として体裁が整っているか、

執筆要領にしたがっているか、

簡潔明瞭に記述されているか、

実証的なデータは適切に位置づけられているか、

注や参考文献は本文と対応しているか、

専門用語の使用は適切か、

語句や文法的な誤りがないか、

誤字脱字はないか、

句読点に誤りはないか、

英文要約などの英文表現は適切か、(必ずしも和文要約の直訳である必要はなく、英文としてまとまっていること)

字数は規程に従っているか、

など、内容及び形式に関する留意点があります。

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

5. 査読にあたっての判断事例

(1) 完成度において不十分であるが掲載を考慮できる場合

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

検証は十分とはいえないが、理論や定式化が学問の発展に有用である。

考察は十分とはいえないが、新たな理論の形成・促進に有用である。

文献調査は十分とはいえないが、研究の位置づけは明確である。

比較研究は十分とはいえないが、適用例としては意義がある。

考察は十分とはいえないが、社会的、または、歴史的に重要な事例の評価として意義がある。

考察は十分とはいえないが、特定の社会活動の評価として意義がある。

論文の構成や表現は適切とはいえないが、内容は評価できるものがある。

論理性は十分とはいえないが、実務上の有用性がある。

有意義な実践・調査報告である。

(2) 掲載を考慮するのが困難と判断される事例

問題意識や問題の設定が不明確

基本的な用語の概念の理解や分析枠組が不明確または不適切

論拠とするデータ等の信頼性が乏しい

論旨の明確さや論証の適切さが無い

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。

～ 投稿案内 ～

日本評価学会では、「日本評価研究」掲載のための投稿原稿を募集しております。1年に2回（3月末/9月末）の投稿締切を設けておりますので、ご興味をお持ちの方は投稿規定・執筆要領をご参照のうえ、奮ってご投稿ください。

投稿先：特定非営利活動法人 日本評価学会事務局
〒135-0047 東京都江東区富岡2-9-11京福ビル
財団法人国際開発センター内
TEL: 03-3630-8031 FAX: 03-3630-8095

『日本評価研究』第6巻第1号

2006年3月27日

編集・発行 特定非営利活動法人 日本評価学会
〒135-0047 東京都江東区富岡2-9-11京福ビル
財団法人国際開発センター内
TEL : 03-3630-8031 FAX: 03-3630-8095

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。

The Japanese Journal of Evaluation Studies

Vol. 6, No. 1, March 2006

CONTENTS

Special Issue: Evidence-based Evaluation

- Editor's Note Ryo Sasaki
- The Genealogy and Direction of Evidence-based Medicine (EBM):
The Future and Role of the Cochrane Collaboration to Achieve Health Care Evaluation
Tomoya Masaki, Kiichiro Tsutani
- Utilization of Evidence in Criminal Justice with a Focus on Prevention and Treatment of
Delinquency Hiroshi Tsutomi
- Evidence-Based Education
What Works Clearinghouse in the United States Satoko Tanabe
- Challenge of Evidence-Based Evaluation in ODA Sector: Poverty Action Lab Ryo Sasaki

Articles

- Evaluation of Social Capacity for Urban Air Quality Management
Kazuma Murakami, Shunji Matsuoka
- The Influence of the Decentralization in the Area of Junior Secondary Education in Indonesia
An Empirical Study Using Causality Model on the Relationship between Educational
Environment and Outcome Atsuko Toi, Hiromitsu Muta
- The Twentieth-Year Review of Japanese Structural Adjustment Loans:
Political Background and Decision-Making of the Japanese Government Yayoi Tanaka
- A Quantitative Evaluation Utilizing Standardized Effect Unit: Application to the Evaluation
of Foreign Student Policy and Regional Cooperation Program Yuriko Sato
- The Evaluation and Analysis of a Policy about Information Flowing
The Case of Eco-Mark Program Takashi Hirahara
- An Analysis of Diffusion and Developmental Process in Local Governments' Evaluation and
Performance Measurement Shun'ichi Furukawa, Harumi Morikawa

Research Note

- A Study of Evaluation Methodology for Donor-funded School Construction Projects in
Developing Countries through Establishing Evaluation Guidelines
Satoshi Morita, Yasuyuki Sagara

Development

- A Study of Student Intern Program for Its Promotion:
The Case of Evaluation on Nonprofit Organization Operation of "Action with Lao Children"
(DeknoyLao) Isamu Moteki

Report of 6th Annual Conference of the Japan Evaluation Society