

日本評価研究

The Japanese Journal of Evaluation Studies

Vol. 7, No. 1, March 2007

特集：教育評価

特集「教育評価」について 牟田 博光

学校評価の理論と実践の課題 長尾 眞文

自己評価の活用による学校改善の実践報告
石田 謙豪 平 恵津子 住元 しのぶ 長尾 眞文

大学評価・学位授与機構による試行的大学評価事業の評価
齋藤 貴浩 林 隆之

開発途上国における教育評価に関する理論的比較研究
- 国際学力調査、学校調査、世帯調査の視点 - 西村 幹子

教育分野における参加型開発支援プロジェクトの評価に関する基礎的考察
石田 洋子

ノンフォーマル教育援助における参加型評価手法の活用
- 「利害関係者が評価過程に評価主体として関わること」の意義 - 源 由理子

初等教育分野におけるプログラム・ベースド・アプローチの援助に関する評価分析
ホンジュラス及びベトナムのEFA-FTIの事例から 三輪 徳子

研究論文

段階評価における項目の重み導入による三次元的分析の試み
- 大学における授業評価分析を事例として -
星野 敦子 北原 俊一 新行内 康慈 安達 一寿
綿井 雅康 牟田 博光

Economic Burden of Disease and Cost Benefit Analysis of Emergency Medical
Service (EMS) System in Romania
Takao Ozaki, Marius Enescu, Moazzam Ali, Chushi Kuroiwa

社会的環境管理能力の形成プロセスに係る因果構造の分析：
都市大気汚染対策を事例にして 村上 一真 松岡 俊二 金原 達夫

政策評価の客観性をめぐる一考察
- 不良債権処理に伴う雇用対策評価をめぐって - 西本 哲也

研究ノート

Trends in Air Pollutant Concentrations and Related Environmental Legislation in India
Miyuki Nagashima

多様化する学士課程教育の自己評価方法 串本 剛

『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

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(三菱UFJリサーチ&コンサルティング(株))
Yoshiaki OKAMOTO

小野 達也(鳥取大学) 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 眞理子(筑波大学)
Ryo SASAKI Mariko SATO

渋谷 和久(国土交通省九州地方整備局) 鈴木 絲子(神戸女子大学)
Kazuhisa SHIBUYA Itoko SUZUKI

田中 弥生(大学評価・学位授与機構) 松岡 俊二(早稲田大学[2007年4月~])
Yayoi TANAKA Shunji MATSUOKA

村松 安子(東京女子大学)
Yasuko MURAMATSU

事務局
Office

〒140-0002 東京都品川区東品川4-12-6 日立ソフトタワーB 22F
財団法人国際開発センター内
特定非営利活動法人日本評価学会
編集事務担当 藪田 みちる
TEL: 03-6718-5931, FAX: 03-6718-1651
E-mail: jes_info@idcj.or.jp

日本評価研究

第7巻 第1号 2007年3月

目次

特集：教育評価

牟田 博光

特集「教育評価」について.....1

長尾 眞文

学校評価の理論と実践の課題.....3

石田 謙豪 平 恵津子 住元 しのぶ 長尾 眞文

自己評価の活用による学校改善の実践報告21

齋藤 貴浩 林 隆之

大学評価・学位授与機構による試行的大学評価事業の評価33

西村 幹子

開発途上国における教育評価に関する理論的比較研究

国際学力調査、学校調査、世帯調査の視点47

石田 洋子

教育分野における参加型開発支援プロジェクトの評価に関する基礎的考察61

源 由理子

ノンフォーマル教育援助における参加型評価手法の活用

「利害関係者が評価過程に評価主体として関わること」の意義73

三輪 徳子

初等教育分野におけるプログラム・ベースド・アプローチの援助に関する評価分析

ホンジュラス及びベトナムのEFA-FTIの事例から87

研究論文

星野 敦子 北原 俊一 新行内 康慈 安達 一寿 綿井 雅康 牟田 博光

段階評価における項目の重み導入による三次元的分析の試み

大学における授業評価分析を事例として105

Takao Ozaki, Marius Enescu, Moazzam Ali, Chushi Kuroiwa Economic Burden of Disease and Cost Benefit Analysis of Emergency Medical Service (EMS) System in Romania	117
村上一真 松岡俊二 金原達夫 社会的環境管理能力の形成プロセスに係る因果構造の分析： 都市大気汚染対策を事例にして	131
西本 哲也 政策評価の客観性をめぐる一考察 不良債権処理に伴う雇用対策評価をめぐって	147
研究ノート Miyuki Nagashima Trends in Air Pollutant Concentrations and Related Environmental Legislation in India	161
串本 剛 多様化する学士課程教育の自己評価方法	171
第7回全国大会開催報告 第7回全国大会プログラム（実績）.....	185
共通論題セッション報告	189
自由論題セッション報告	192
委員会活動 企画委員会	200
研修委員会	201
広報委員会	202
国際交流委員会	203
春季第4回全国大会のご案内	204
特定非営利活動法人日本評価学会定款	205
日本評価学会誌刊行規定	215
日本評価研究投稿規定	217
日本評価研究執筆要領	219
日本評価研究査読要領	222
英文学会誌の刊行について	224
Publication Policy of the Japanese Journal of Evaluation Studies	225
Information for Contributors (For English Papers).....	227
Writing Manual of the Japanese Journal Evaluation Studies (For English Papers)	229
Referee-Reading Guideline	231

【巻頭言】

特集「教育評価」について

牟田 博光
東京工業大学

国内では、中央教育審議会の「地方教育行政」に関する答申（1998）で、学校が地域に開かれ、信頼される組織となるように、自己評価の実施と結果の公表を提案したことを手始めに、2002年に学校設置基準が改訂され、自己評価が努力規定とされ、全国的に学校評価システムの導入が進められてきた。2006年には文部科学省により「学校評価ガイドライン」が定められ、各学校は自己評価、外部評価を実施しその結果を公表することが提示された。大学に関しては、1991年の大学設置基準の改定により自己評価が努力規定とされ、次いで義務化された。さらに、第三者評価の必要性が説かれ、2002年には義務化された。2004年に国立大学が法人化された後は、中期目標、中期計画を作り、国立大学法人評価委員会が評価することとなった。このように、学校・大学の自己評価が推奨され、次いで義務化され、第三者評価が推奨され、次いで義務化されるといったトップダウンの経緯をたどっている。学校・大学が全国一斉に評価の網が法的にかけられたところから、表面的な評価実施率は高いものの、現場の教職員の評価に関する知識はきわめて限られたものでしかなく、大きな混乱や反発があるのが実態であり、評価をどのようにして教育行政に円滑になじませ、PDCAサイクルに真に生かしていくかが問題となっている。わが国の教育行政に即した、学校評価理論の構築と評価実践のベストプラクティスの蓄積が必要とされている。

長尾論文は、学校評価の理論を学校評価の仕組みの理論と学校評価の実践の理論の二層構造の枠組みを示している。石田（謙豪）他論文は、この理論に基づく公立小学校による学校評価の実践報告である。また、齋藤・林論文は第三者評価機関である大学評価・学位授与機構が2000年度から2003年度まで国公立大学等延べ550組織に及び試行的評価を実施した評価結果を分析し、評価設計、評価結果の妥当性、その成果について論じている。

一方、開発援助の世界では、教育分野、とりわけ、ベーシック・ヒューマン・ニーズを重視する関係もあり、低所得国での基礎教育や学校外のノンフォーマル教育部門への援助が増加しているが、自律発展性をどう担保するかが大きな問題となっている。学校建設、教員や地方教育行政官の研修など、効果的な学校や教育行政の仕組みを作る段階での評価を行わなければならないが、ある程度できあがっている教育の仕組みをどう改善していくかということにまで至っていない場合が多い。自律発展性を担保する解決策の一つとして、様々な形の参加型援助やプログラム援助の効果が期待されている。また評価を事業改善、能力開発の一環と位置づけ、参加型で評価する試みも行われている。

西村論文は、教育評価についての既存の国際学力調査、学校調査、世帯調査を比較・対比させ、開発途上国においては各種調査の結果が大きく乖離し、総合の必要性が特に大きく、定量的なデータだけでは得られない推論を定性的な調査から得ることの重要性を示している。石田（洋子）論文は、限られた資源で教育改善を進めるために、教員配置や教科書配布、学校施設維持管理への技術指導など、住民の活動に対する行政サービスを行う教育行政官の能力向上の必要性について述べている。源論文は、定型化が困難なノンフォーマル教育において、参加型評価は、利害関係者間の「対話」を通して、態度・行動変容といった質的側面の指標化を可能とし、より適切な提言策定とその活用につながる可能性が高く、関係者の評価能力向上を含めたマネジメント能力の向上、ひいては組織強化につながる可能性があることを示している。三輪論文は、ホンジュラス及びベトナムの「万人のための教育 ファスト・トラック・イニシアチブ」を事例に、開発成果実現の観点から、プログラム・ベースド・アプローチによる援助の形成評価を行っている。

【総説：依頼原稿】

学校評価の理論と実践の課題

長尾 眞文

広島大学

nagaom@hiroshima-u.ac.jp

要 約

1998年（平成10年）の中央教育審議会の「地方教育行政」に関する答申で、学校が地域に開かれ、信頼される組織となるように、自己評価の実施と結果の公表を提案して以来、全国的に公立学校による学校評価システムの導入が進められてきた。現在では公立小中学校による自己評価の普及率は95%を超えるまでになったが、各学校で実践に当たっている教職員の評価に関する知識はきわめて限られたものでしかない。また学校評価に関する文献は近年増加の一途を辿っているが、その多くは実践の後追的な内容で、理論的な整理の試みは余り例がない。本稿では、学校評価の理論を学校評価の仕組み（学校評価システム）の理論と学校評価の実践（評価手法）の理論の二層構造から成るものと捉え、日本における実践を念頭においてその枠組みを提示することを試みる。結びでは、この枠組みを実践に移していく上での課題の整理を行っている。

キーワード

評価理論、学校評価システム、自己評価、外部評価、メタ評価

1. 学校評価の普及と理論的課題

近年国内の公立学校で学校評価の取組が急速に普及している。文部科学省（文科省）初等中等教育局（2006）の調査によれば、2004年度には公立学校の96.5%が教職員による自己評価を、78.4%が保護者・地域などによる外部評価を実施していることになっている。しかしこの高い普及率を額面通りに受けとめて、学校評価が定着していると考えるのは正しくないようである。例えば上記の調査によると、自己評価の結果を公表している学校の比率は42.8%に過ぎず、半数以上の学校につ

いて単なる学校の内部的な反省や振り返りを自己評価と称しているのではないかとの疑義が生じる。また、最近増えてきた学校経営や学校評価に関する著作の多くで、「新しい」学校評価の実践の必要を説いているが（例えば、木岡 2004、西村・天笠・堀井 2004、金子 2005）、それは取りも直さず現在行われている学校評価の多くが単に見かけを取り繕う評価であったり教育委員会にやらされている評価であることの裏返しであるようにもとれる。

一方、学校評価に関する関心の高まりが評価手法も含めてそれに関連する評価理論の理解の深化

を誘引しているかというところではない。この分野の研究の第一人者とも言うべき木岡（2003）によれば、これまでの学校評価研究は「あるべき学校評価」論に基づく実践試案やチェック・リストの作成に終始し、その普及や定着を阻んでいる要因や有効な評価の取組に関する体系的な理解の創出に役立っていないという。実際に学校評価の日本語文献で学校評価の「理論」に言及しているものは極端に少ない。しかし、実際に評価を実施している教職員の中には、学校評価の理論的理解の必要を感じている者が少なくない。筆者は、2003年から広島県立教育センターが毎夏県内の公立小・中・高等学校の教員20余名を対象として開設する学校評価研修講座に講師として参加しているが、受講者の参加希望理由の第一は常に学校評価の理論に関する理解を深めることである。

理論的な裏打ちのない評価の実践は、学校評価に限ったことではない。政策評価や行政評価といった組織の評価も同様で、説明責任（答責性）を突きつけられた公的機関が待たなしで実践的取組を迫られ、評価理論の構築を図る余裕もないのが現状である。ここで評価理論とは、Shadish, Cook & Leviton（1995）の言う評価方法の選択の理論のことで、その基本は個別の事業の評価をいかに行うかの方法論から成っている。事業の評価は、組織の評価と異なり評価の対象が時間的・現象的により限定されるため、事業計画の理論（事業の想定している因果関係をどう捉えるか）、評価結果の活用の理論（評価を誰のために行うか）、判断に関わる価値の理論（判断基準を何に求めるか）といった構成要素を組み合わせることで系統的に扱うことができる。組織の評価は、組織が行う個別事業の評価の総和の視点に加えて組織自体の存在価値や存続の必要性の視点も備える必要がある。それでも、Scriven（1991）は、評価の理論を評価手法の理論と評価の社会的役割の理論から成るものとするにより、より幅広い事業、組織の双方を包含する理論の概念を提示する。

学校評価の理論も、学校評価の仕組みに係わる理論（評価システムの理論）と学校評価の実施・作業に係わる理論（評価手法の理論）の二層構造から成ると考えられる。学校評価システムの理論

は、主として学校が教育行政の仕組みの中でどのように位置づけられているかといった教育制度のあり方に依存する。学校評価手法の理論は、より実践的な内容を含み、評価の目的、範囲、資源、活用方法等に依存する。しかし、評価の仕組みと実施の理論的枠組みが定まっても、評価の実践はまた別である。それ故、評価は「科学プラス技能（science and art）」あるいは「科学プラス技能プラス技巧（science, art and craft）」と言われる（Conner 2001；Compton, Baizerman & Stockdill 2002）。評価の実践的課題を検討することは、同時に評価理論の限界を示すことでもある。

以上の評価理論に関する理解を踏まえて、本稿では、まず国内で急速に普及しつつある学校評価を理論的に把握するための枠組みを、学校評価の仕組みと学校評価の実施方法の順に提示する。最後に、その理論的な理解に基づく評価慣行が普及・定着するための条件あるいは課題について検討することにする。

2. 学校評価の仕組み：評価システムの理論

（1）教育行政制度と評価の仕組み

学校をどう評価するかは、誰が何の目的で評価を行うかによって決まるが、それは第一義的に学校がどのような教育行政制度の中に置かれているかに依拠している。国家が集権的に教育を管理し、それを徹底するための教育行政の仕組みを作り上げている場合には、学校評価はその仕組みを機能させるツールとして活用されるであろう。その反対に、教育が分権的に行われ、学校が教育の内容や組織経営のあり方にかかなりの裁量権を有するような教育行政下では、学校評価の仕組みはより自律的なものとなるであろう。実際には、その両極端の間にほとんどの国の教育行政制度があって、学校評価の仕組みも多様な形で出来上がっている。学校評価の仕組みに焦点を絞って理論的な枠組みを考えるには、誰が評価するか、何の目的で評価するかを分けて考える必要がある。

まず誰が評価するかであるが、概念的に内部評価と外部評価に分けて考えられる（Scriven 1991）。

内部評価は組織が内部的に行う評価、外部評価は組織外の評価者が行う評価である。内部評価にも組織構成員が全員参加する形で行う自己評価と組織内部に評価部局を設けて行う内部専任評価がある。外部評価にも組織が組織外の評価者を選んで行う外部（委託）評価と組織の意思とは無関係に外部の個人・団体が組織を評価する第三者評価がある。これらは概念的な区分であって、実際には境界はあいまいで、多様に組み合わせられて評価が行われることが多い。評価慣行の普及している欧米諸国では、伝統的に外部評価が主体であるが、1980年代後半から組織改善に対する関心の高まりを反映して内部評価が重要視されるようになった（Love 1991）。学校評価についても同様で、欧米諸国では長い間外部評価が中心的役割を担ってきたが、近年内部評価、特に学校自己評価が注目を集めている（Nevo 2002）。

次に何の目的で評価するかであるが、Patton（1997）によると、評価には結果の使い道によって「判断をくだすため」、「改善を促進するため」、「知識を創出するため」の3つの目的が考えられる。「判断をくだすため」の評価は、総括的、監査的に行うもので、説明責任（答責性）を問う評価がその典型である。「改善を促進するため」の評価は、形成的に行うもので、継続的な組織改善のツールとして使われる。「知識を創出するため」の評価は、実践的と言うより学術的で、ある事象や活動に関する一般的理解が目的である。学校評価については、通常「説明責任（答責性）」と「改善」が目的として挙げられる。しかし、国が教育の現状や課題について政策資料を得るためにランダムに選んだ学校で学力テストを行ったり、診断的評価を行ったりして知識創出目的の評価を活用することもある。

学校評価の仕組みは、内部評価・外部評価の評価主体軸と説明責任・改善の評価目的軸の組み合わせによって決まる。国家管理による教育を行っている国では、説明責任目的で外部評価を中心とする監査的な学校評価システムを構築するであろうし、各学校が自らの管理運営組織を持つような教育体制下では、改善目的で内部評価と外部委託評価を組み合わせたシステムが主体となろう。教育行政制度のあり方によって、学校評価システム

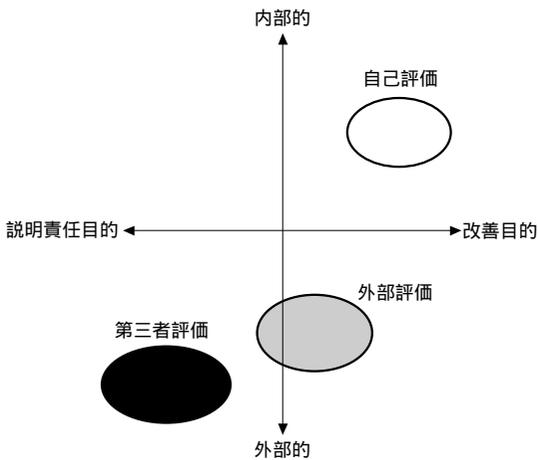
の機能の仕方も異なり、国レベル、地域（地方教育委員会）レベル、各学校レベルで多様なシステムへの関与の形が考えられる。

学校評価システムの機能の仕方をどう評価するかは、それ自体難しい問題である。当然評価システムの目的を達成できているかどうかの有効性の視点、それが無理や無駄がなく実現できているかどうかの効率性・実行可能性の視点は欠かせないであろう。また順法性、合規性、道徳性を含む正当性の視点も重要である。さらには、それらの視点を総合してシステムに対する信頼性が備わらなければ仕組みができたことにならない。そして信頼性を左右する重要な要因として、評価システムが欠陥を認めた学校を対象とする支援システムに連動しているかどうかが挙げられる。もちろん評価の結果、統廃合の形で淘汰される学校も出てくることは予想されるが、ほとんどの学校が抱える何らかの弱み、欠陥、課題に対して、改善の取り組みを後押しする機能が備わっていなければ、システムの有効性は期待できず、正当性も怪しいこととなる。多様な評価の視点がどう作用するかは、実際に評価システムが置かれている状況により異なる。

（2）学校内部評価と外部評価：欧米諸国の場合

欧米諸国の学校評価システムには多様な形が見られる。学校評価の慣行の普及している米国を例にとると、分権・分散的な教育行政制度の下で、教育の質の保証・改善のための機関による認証評価、州政府の法令により実施する監査的評価、学区単位で行う内部専任の評価、教育分野の民間団体が調査目的で行う第三者評価、学校自身の改善目的評価が様々に、並行的に行われている（Sanders et. al. 1995）。教育行政的には全体が統一的な体系を持っている訳ではないが、学校評価の取り組みがそれぞれに機能を発揮している。これを学校の立場から見た内部・外部評価の別、評価の目的の別の2軸で捉えると図1のように表現できるであろう。ここで各評価の濃淡は、その相対的影響力を表す。学校評価システムの中心は、一方で教育予算を握る州政府及び関連組織が行う第三者評価と他方で学校が認証組織との契約に基づい

図1 学校評価システムの構造：米国の場合



(出所)筆者作成

て行う外部評価である。前者は学校が望ましい水準の教育を提供しているか否かを共通学力テストによって検証すること等により監査的評価を行う³。後者については前提として各学校が自己評価を行うことになっているが、認証の目的は学校が継続的な改善を心がけるように仕向けることにあり、かなりの専門性を備えたスタッフが具体的な改善策等について指導・助言する⁴。但し、その対象は認証申請をする学校に限られる。米国において学校評価システムが実際に学校教育の改革、学校経営の改善にもたらす効果は地域によって異なり、その格差が問題になっている(中留 1999)。

欧米諸国の学校評価システムの特徴は、各学校が自律的に改善を志向するように仕向ける組み立てにあり、それを機能させているのが教育行政組織による第三者評価、認証機関等による外部評価ということになる⁵。教育行政組織による第三者評価は、結果が予算配分に直結している場合が多く、成果ベース管理の風潮に乗って一般的支持は受けているが、学校は必ずしも歓迎していない。学校の改善には継続的な努力が必要で、それにはそのような努力の総括評価を行うだけでなく、学校自身による形成的評価を支援すべきとの考えである。

最近では、自己評価の役割をさらに強調して、評価を行うこと自体が学校改善であるという継続

的なモニタリングに基づく実践論も主張されている。この考え方は、MacBeath & McGlynn (2002) が「開発型」と呼ぶ視点で、教育活動の成果が圧倒的に教員の資質、姿勢、努力に影響されることを踏まえて、評価が教員の教育実践を支援すべきとする。Hopkins & West (2002) の英国における研究によると、「学校改善としての評価」の考え方は、次のような条件下で特に有効なようである：(i) 学校が改善に主体性を持つこと、(ii) 説明責任の圧迫がないこと、(iii) 系統的・意図的な方法の採用、(iv) 選択的課題への集中、(v) 改善プロセスと関係者の役割の明確化、(vi) 評価の改善計画における明確な位置づけ、(vii) 外部支援の適宜の活用。個々の学校で、評価の習熟度が上がるにつれてこのような取組が一般化することも十分考えられる。しかし、学校評価システムを個別の学校単位でなく、公教育全体のレベルに戻して考えると、日本におけると同様に欧米諸国においても「外部評価は嫌われ、内部評価は信用されない」という状況は変わらず、結局両者の適当な組み合わせの実現が最善ということになる(Nevo 2002)。

(3) 学校自己評価と外部評価：日本の場合

日本では一部の地域や学校による学校評価の取組の試みは戦後間もない頃からあったが、学校評価が全国的規模で取り上げられるようになったのは最近のことである⁶。その背景には、1998年9月の文科省中央教育審議会(中教審)の「地方教育行政」に関する答申に端を発する、国からの学校自己評価推進の動きがあった。公立学校が地域に開かれ、信頼される教育の場となるように、学校は毎年自己評価を行い、結果を公表すべきとの方針は、2002年に文部科学省令による学校設置基準の改定で努力義務が課せられた。2005年10月の中教審答申「新しい時代の義務教育を創造する」では、義務教育の構造改革の方向として、国が教育目標の設定とその実現のための基盤整備について責任を持ち、実施過程を担う市区町村や学校の権限と責任を拡大して、その自主性・自律性を強化し、「義務教育の成果を検証する仕組みを国の責任で整備し、教育の質が保証される教育システムへの転換を図ること」が提示され、学校評

価が検証のための主要な手段であることが明示された。そして、2006年3月には、文科省が学校評価ガイドラインを作成し、各学校が自己評価と外部評価を実施し、その結果を説明・公表すること、学校設置者等が評価結果に応じて学校に対する支援や条件整備等の措置を講じることを目指すべきことを、具体的な評価方法と評価項目の例示も含めて提示した⁷⁾。

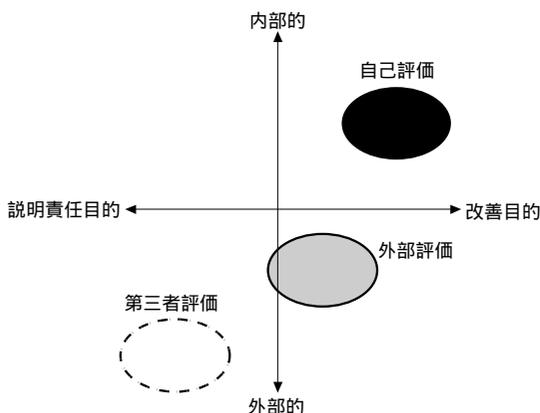
以上に明らかなように、日本における学校評価の動きは中央からのトップ・ダウンの圧力によって展開している。図2に示すように、その中核は各学校が内部的に改善目的で行う自己評価であり、自己評価の信頼性、透明性を高めるために外部評価をあわせて行うことが想定されている。冒頭で述べたように自己評価を実際に行っているほとんどの教職員は、学校の組織評価に関する理解や経験を持っているわけではなく、多くの学校では未だ先事例に模してとにかく評価作業実施の事実を重ねているに過ぎない。最近各地の教育委員会で頻繁に開かれている学校評価のセミナーやワークショップでも、専らどう評価するかの実験の交換が主で、評価の実施がいかに役立ったかに踏み込んだ議論は余り聞かれぬ。外部評価は、前述の文科省のガイドラインが示唆するように、多くの場合学校あるいは教育委員会が地域や保護者の代表に委嘱して、学校訪問・授業や行事の観察、学校管理者・教職員からの聞き取り等に基づ

く自己評価結果の二次評価を行っている。そこでの問題は、外部評価者としての専門性を備えた人材が地域に少ないこと、居たとしても時間の確保が困難なこと、予算不足で報酬を支払うことが出来ないことで、結局は町内会会長、民生委員、退職校長といった学校に「近い」人物になり、外部評価に期待される客観性や信頼性醸成効果が期待できないことである⁸⁾。

学校による自己評価と外部評価の組み合わせが期待した効果を創出していない状況で、文科省は前述した国の責任で教育の成果の検証を行う仕組みを作る方針から、省内に外部専門家による学校評価調査研究協力者会議（筆者も委員として参加している）を設けて第三者評価のあり方を検討している。2006年度には、学校評価の現状と課題の確認や外国の第三者評価制度についての検討に加えて、各都道府県および政令指定都市で小中学校をそれぞれ1校ずつ計124校選んで第三者評価の試行も行っている（文部科学省 2006）。2007年度にはさらに試行の規模を拡大して実施するとともに、第三者評価制度の最終的な形についての提言をする予定である。そこでの主要な課題は、国が行う第三者評価の目的をどう設定するかである。試行のモデルとも考えられる英国の制度では、教育の質を保証するために設けられている教育水準局（OFSTED）が、全国の公立学校を対象に数年間に1度学校視察団を派遣して監査的な評価を行い、各学校の教育・経営改善を促すとともに、問題のある学校については教育指導部による直接的関与でフォローする仕組みが出来ている⁹⁾。日本の場合、国が行う第三者評価にそのような制度的機能を付与するには、新たな仕組みづくりのために現在の教育行政制度の大幅な変革（例えば、教育委員会が学校設置者として有する学校管理機能をどうするか）やかなりの財政的・人材的措置が必要となり、その実現には高度の政治的判断が必要となる。より現実的な対応として、フランスで行っているように、毎年100校程度の公立学校をランダムに選んで健康診断的な第三者評価を行い、その結果を政策形成に資料として活用することが考えられる。

現在の日本の教育行政制度を前提とすると、全国的な学校評価システムの構築に県市町村レベル

図2 学校評価システムの構造：日本の場合



(出所) 筆者作成

の教育委員会が果たす役割を外すことはできない。特にシステム全体の目的が地域に開かれ、信頼される学校づくりであるからなおさらである。実際に各学校による自己評価・外部評価の仕組みの構築には教育委員会が重要な役割を果たしている。例えば、広島県教育委員会を例にとると、学校評価システム検討会議の設置による『広島県における学校評価システムの在り方』ガイドラインの作成と公表（2002年度）、モデル校による実践研究に基づく『学校評価資料 - 協力校の実践事例』の刊行（2002年度）、学校設置基準の改正による自己評価の義務規定化（2003年度）、県立教育センターによる小中高の教員を対象とする学校評価（4日間）研修講座の開催（2003年度以降毎年）といった具合に、各学校による評価システムの確立を継続的に支援してきている¹⁰。教育委員会には、学校設置者の立場から学校経営の改善を指導する責任があり、学校に対して管理主事・指導主事による指導・助言を経常的に行うことになっている。学校の自主的・自律的な学校改善努力を支援するため、地域の大学や研究機関等の支援も巻き込みながらの教育委員会による模索が全国各地で進められている（長尾・和佐）。さらに、広島市教育委員会では、市独自の第三者評価制度を設けて、課題を抱えている学校の支援の方法を模索しているが¹¹、全国的に形成される第三者評価の仕組みも、何らかの形で教育委員会が関与するものと思われる。

3. 学校評価の実施：評価手法の理論¹²

(1) 学校経営と学校評価

学校評価システムが決まると、次に各学校で実際に評価をどのように行うかの方法が問題となる。文科省の指導により、学校は教育活動を含む学校運営に「計画（Plan） - 実行（Do） - 評価（Check） - 改善（Action）」の経営サイクルの考え方を導入してきた。このサイクルの中に学校評価を位置づけることが、評価実施の第1歩である。学校は、表1にあるように、教育目標の確認から出発して、学校の内部的状況と外部条件の検討に

よるニーズ分析を行い、それに基づきミッション・ビジョンを策定、ビジョンを具体化して学校経営の中期目標（3年）と短期目標（1年）を特定した後、目標達成のための手段・活動を明確化することで、学校経営計画が出来上がる。学校評価計画は、経営目標とその達成のための手段・取組の間の因果関係を確認することから出発し、計画の進捗状況の把握も含めて評価の方法を確定し、外部評価の実施予定も含む年間評価活動スケジュールを作成することで出来る。評価の枠組みに当てはめて計画を検討することで、計画の実行可能性のチェックができる。

表1 学校経営計画から学校評価計画へ

<p>教育目標の確認</p> <p>ニーズ分析（学校の「強み」、「弱み」、「有利化/不利化要因」）学校状況の把握・分析/学校を取りまく外部要因の把握</p> <p>ニーズ分析からミッション（使命）の決定へ 教育の使命は？学校教育の役割は？ 自校では、何を指して教育活動を進めるか？</p> <p>ミッションからビジョン（未来像）へ 使命の先にある自校の将来像は？ 社会・地域・学校関係者との関係は？</p> <p>ビジョンを具体化して経営目標に絞り込む 中期経営目標（3年程度の達成目標） 短期経営目標（1年単位の達成目標） 目標達成手段・活動の明確化による実施計画の作成</p> <hr/> <p>学校評価計画の作成</p> <ul style="list-style-type: none"> ・ 評価の枠組み（評価すべき目標と手段/取組及び評価方法の明確化） ・ 計画実施の進行管理（モニタリングの方法） ・ 外部評価実施計画 ・ 年間評価実施スケジュール

（出所）筆者作成

(2) 学校経営目標と評価方法

学校経営目標には、大別して学校組織の全般に係わる全方位型とよりの絞った特定課題指向型の二つのタイプがある。評価の視点から見ると、全方位型は組織評価、特定課題指向型は事業評価の性格を有する。表2に示すように、経営目標の

表2 学校経営目標と評価方法

経営目標タイプ	評価目的	評価方法
全方位型 組織評価	説明責任（答責性）の徹底・共通理解形成	学校経営実績の全体的（全分掌における）把握・提示
	パフォーマンスの改善	長所・短所の発見、対応の検討と次年度計画への反映
特定課題志向型 事業評価	学校の改善（学校の特色づくり）	課題に対応するための施策の実施プロセスと達成状況の把握、改善策の提示
	問題解決（学校の緊急課題の解決）	学校が緊急課題への対応のために講じた施策・措置の有効性、妥当性、効率性の検討と場合によっては対応の継続のための施策の提示

（出所）筆者作成

タイプによって評価の目的や方法は異なる。全方位型の評価では、説明責任の徹底や学校の全体的パフォーマンスの改善が目的となり、それぞれ全分掌における実績の把握や長所・短所の発見が方法としてとられる。一方、特定課題志向型の評価では、学校の特色づくりのような絞った改善や学校が抱える緊急課題の問題解決が目的となり、それぞれ具体的な対応策が機能するかどうかを検証する方法が選ばれる。

（3）学校評価の枠組みづくり

成果目標と努力目標の区別の必要

学校評価の枠組みを作る時に重要なことは、学校経営計画を作成する段階で、学校が目指す成果目標（例えば、生徒の学力向上）とその達成のための努力・取組目標（例えば、繰り返し学習、教員研修）が明確に区別されていることを確認することである。両者の間には因果関係が想定されている筈で、目標の連鎖体系が成立していなければならない。現在行われている学校評価では、学校・教職員の努力や取組の達成度（研究授業を何回実施した、休憩時間中の屋外活動参加者が増えた等）のチェックに終始していて、それがどのような成果をもたらしたのか（例えば、生徒の理解は増したのか、生徒の体力は向上したのか）という肝腎な設問が忘れられている場合が多い。それ

では、説明責任の遂行にも改善努力にもつながらない。簡潔に表現すると、成果目標は生徒に起きる教育効果の特定化であり、努力・取組目標はその効果を実現するための学校・教職員の措置・施策の設定である。但し、教員の研修の取組が生徒に効果を及ぼすにはある程度の期間を要するので、年度計画では中間的な成果である教員の指導力の向上を挙げることが適当であろう。

学校が掲げる教育・経営目標には、「生きる力」、「豊かな心」、「地域に開かれた学校」のように、数量化・指標化が難しく、具体的な成果目標を設定しにくい場合も多い。その場合は、学校および関係者が合議の上、例えば今年は「生きる力」を体力・持久力で表現するとか、「豊かな心」はあいさつ習慣の習得で捉えるというように、便宜的に成果目標を設定すれば良い。評価する段階では、単にあいさつの回数を数えて量的に表現するのではなく、態度・意識がどう変わったかに関する教師の観察や生徒の作文中の表現で質的側面もあわせて考えることにより、より正確に達成成果を表現することができる。参考までに、「確かな学力」の定着について成果・努力目標の立て方、評価項目、評価指標を例示すると表3の通りである。

ロジック・モデルの活用

成果目標と努力目標の区別ができたところで、評価の枠組みづくりにロジック・モデルを活用す

表3 「確かな学力」の定着

	経営目標	評価項目	評価指標
成果	生徒の学力の向上	表現力、読解力、総合学力の変化	テスト、プロジェクト、作文、発表
	授業の改善	教員の指導力の変化	生徒の授業理解度・満足度
努力	生徒の学習習慣・態度の改善	読書習慣の定着 家庭学習の定着	頻度、読書量、参加者 宿題、保護者観察
	教員の研修	校外研修への参加 校内授業研究の実施	頻度、結果活用度 頻度、内容
	全学体制の強化	シラバス作成の徹底 Team Teaching 授業の活用	達成度、内容 実施頻度、内容

(出所)筆者作成

ることができる。ロジック・モデルは、特定の課題に対する取組が理論上どう組み立てられ、最終成果がどのように実現するかを簡潔に示すための枠組みで、フローチャート、表、図等で表される¹³。学校経営計画でミッション・ビジョンから導かれた中期・短期目標から評価すべき主要課題を選択し、それぞれについて 取組・活動、必要な投入、活動結果、成果・インパクトを体系的に

表現してみるものである。表4は、分かりやすい授業の実現のための取組について例示したものである。

評価データの収集方法

評価結果が信頼を獲得するには確かな根拠に支えられなければならない。しかし、不必要に大量のデータを集めることは避けなければならない。それは非効率であるだけでなく、本務の教育で多

表4 ロジックモデルの構成要素と具体例

目標体系	学校のミッション、ビジョン 中期・短期目標	確かな学力を身につけさせる
課題	評価対象課題 課題に関わる目的	分かりやすい授業の実施 生徒の理解力の向上
活動	具体的取り組み・活動	分かりやすい授業ガイドラインの作成 校内研修の実施 公開授業研究の開催
投入	人的、物的、資金的投入	教員の研究・研修時間 教育委員会の指導主事の来校 研修費
結果	活動の結果・産物	分かりやすい授業ガイドライン 校内研修回数・参加教員数・時間 公開授業研究報告
成果	想定成果・インパクト (年度当初と年度末の比較)	教員の授業改善(指導主事による観察) 生徒の授業評価 生徒の学力の変化
外部条件	取り組みの前提となる外部条件	教育委員会の研究支援事業指定

(出所)筆者作成

忙を極めている教員の支援を得られないからである。要は適切なデータ収集にある。それには、複数のデータ収集方法を使って、多様な情報源からデータを集め、それらを適当に組み合わせることにより、過不足ないデータ収集を実現することである。データには、多くの管理データ（出欠数、テスト結果）のように定量的なものや数量で表現できない質を表す定性的なもの（例えば、豊かな心、態度の変化、指導方法等に関する言葉による表現）とがある。それらを組み合わせれば、評価の根拠データはより確かなものとなる。

多面的なデータ源からのデータの収集は成果の判断にとって特に重要である。例えば、習熟度別授業の成果を問うのであれば、アンケートによる生徒の意識の変化の把握、教育委員会指導主事による観察、生徒の（習熟度別）学力の変化等のデータを組み合わせることが考えられる。いずれの場合にも時間軸を適切にとってデータを収集することが大切である。

学校評価の実施事例では、教員、生徒、保護者を対象とするアンケートがデータ収集方法として多用されている。確かにそれは多くの対象から容易にデータを取る手段ではあるが、集計等の時間を考えると、その実施にはかなりの手間を必要とする。実施する場合には、評価に不可欠な項目に質問を絞る必要がある。例えば、生徒の学習規律について評価する際に、家庭学習の様子についてのデータを収集するには、保護者アンケート（「毎日何時間くらい家で勉強するか」等）が有用である。また、教員、保護者、生徒に同様の質問をすることにより、結果の比較が可能となる。アンケート以外にも、生徒・保護者・地域の代表に対するグループ・インタビューや外部専門家による観察も有効なデータ収集手段となり得る。

（４）評価の実施：評価データの分析

年初の学校評価計画の作成時にしっかりとした枠組みを作っておけば、評価の実施は経営計画の進捗状況を管理することになる。ロジック・モデルを活用すれば、実施プロセスのモニタリングや中間・最終成果の評価も同一の枠組みで行うことができる。

通常の学校評価では、評価表で全てを表現しようとする傾向が見られる。しかし、余りに多くの評価情報が表に盛られると、分析結果が分かりにくくなる可能性がある。学校評価における評価表はあくまで総括のための表で、それは全校的・集約的表現である。重点項目について太字や枠囲いで表記するとか、備考欄を使って追加データ（例えば、特記すべき学年別データ、数値化できない要素・要因の説明）を示すとかすべきであろう。

評価表を補完する形で文章表現による評価を加えることも大事である。特に低評価や目標不（未）達成項目の説明は、学校に課された説明責任の遂行に相当するので、丁寧な対応が不可欠である。また、評価には何らかの比較の視点が必要要件であり、評価データの分析には、年度初めと年度末の比較や、国・県平均値との比較、近隣他校との比較等も欠かせない。

（５）費用対効果の視点

日本の教育行政制度の下では学校が予算の裁量権を持たないため、学校は予算の管理はしても運用はしない。会計は事務職員の担当で、通常、教員は教育活動を費用タームで考えることもない。そのせいもあって費用対効果の視点が欠如しがちである。それでは学校の教育活動をロジック・モデルで表す時、投入として何を入れるかであるが、それは教員の時間である。教員の多くがいつも時間に追われているとの意識を持っていることに明らかなように、学校で最も希少な資源は教員の時間である。ロジック・モデル的に考えると、教員の時間の投入・配分が最大の成果を生むような計画にならなければならない。そして、評価データの分析には、教員の時間の投入を費用と考え、それに見合った教育的成果が生じたかの費用対効果の視点が不可欠である。

（６）評価結果の公表から活用へ

評価結果の公表は説明責任の遂行のために行うが、公表することでおしまいではない。前述したように、学校評価システムの導入は、学校が地域に開かれ、信頼される組織となるためである。

従って、何よりもまず評価結果が信頼性を獲得する必要がある。それには、生徒、保護者、地域に学校評価の結果を理解してもらわねばならない。実用重視の立場に立つ評価では、評価結果に関心を持つ層の情報ニーズにいかに対応のかが評価の有用性を決める基準となる(長尾 2004)。従って、結果の提示方法は、当然分かりやすく、しかも関心を喚起するものでなければならない。学校の現状と将来展望に関する学校内外の共通理解の形成から学校改善への関係者の協力の獲得に資するものとなればさらに良い。その意味で、学校は評価結果の公表から活用へと意識の転換を図る必要がある。

(7) 外部評価の実施方法

文部科学省の学校評価ガイドラインによると、「外部評価は教職員と地域住民・保護者が学校運営の現状と課題について共通理解をもち協力することにより、教育活動その他の学校運営の改善が適切に行われるようにすることを目的として実施する」ことになっている。

前述したように、この外部評価から専門的な支援を期待することはできないが、学校の評価活動への参加を通して、学校の運営について信頼性を醸成する効果は期待できる。外部評価の方法も、ほとんどの場合学校の自己評価に対する外部評価委員としての判断(2次評価)の提供という形をとっている。外部評価委員の具体的な仕事は、例えば表5のように表現できる。

表5 外部評価委員の仕事

<p>年度当初に学校側と評価作業内容について合意 年2~3回の外部評価委員会に出席 観察目的で若干の学校行事(公開授業、運動会)に参加 学校が行った自己評価結果について個別判断</p> <ul style="list-style-type: none"> ・各経営目標について:結果は適正/不適正/情報不足/ 「判断回避」 ・全体的取組み・資料・説明が適正(当)かどうか 評価委員会意見の集約に貢献 外部評価結果の公表・報告に参加

(出所)筆者作成

外部評価の活用には、学校、外部評価委員の双

方の側に問題がある。学校側の事情としては、客観性・専門性を備えた人材確保の困難、外部評価者の時間確保の困難、対価支払いに関する予算制約、評価結果への対応における苦勞(無理な提案でも無視できない)がある。一方、外部評価委員の側の事情としては、未経験の作業、学校に関する知識の不足、時間制約、動機の欠如、不明瞭な責任範囲等がある。一部には外部評価の役割を積極的に評価する意見もあるが(西村・天笠・堀井 2004; 小島 2004)、これらの阻害要因の克服は容易ではない。教育委員会の中には、外部評価委員に対して研修を行うことにより、学校評価システム確立に向けての責任を意識してもらおうように図っているところもあるが、未だ欧米諸国のある程度専門性を備えた外部評価委員が期待されている「友好的直言居士(critical friend)」(Doherty et. al. 2001)の域に達するにはかなりの年限を必要とするように思われる。

(8) メタ評価によるチェック

メタ評価とは評価の評価のことで、実施した評価が「良い評価」であるかどうかを評価することである。自己評価にとってメタ評価を行うことは、一種の自浄的な意味がある。Scriven(1991)は、メタ評価は評価を専門とする者にとっては不可避の義務だとし、そのためのチェックリストを提示している。米国ではかなり以前からいくつかの教育機関が合同で策定した評価基準がある(Joint Committee on Standards for Educational Evaluation 1994)。評価基準は、有用性、実施可能性、正当性、正確さの4つで、それに基づく学校評価モデルも開発されている(Sanders et. al. 1995; Stufflebeam 2000)。日本では未だ評価の基準は策定されていない。ここでは、筆者が作成した学校評価のメタ評価チェックシートを文末に参考資料として提示しておく。

4. 結びに代えて: 学校評価の実践の課題

学校評価システムの導入が全国的規模で始まってから未だ日が浅い。公教育の改革・改善の必要

が叫ばれる中で、学校も評価の実施を受け入れざるを得ず、全国各地で実践の積み重ねが始まっている。しかし、その多くは他県、他校の先行事例を模したものであり、評価の理論の理解に基づくものではない。評価の理論はその実践的性格のゆえに、現実世界での評価経験の中から形成されるべきものである。本稿では、筆者の広島県における学校評価システム構築への関与を下敷きとして、内外の学校評価文献を参考に理論的理解の枠組みの構築を試みた。その深化のためには、実践的課題のさらなる検討が不可欠である。以下では、本稿の結びに代えて、既に顕在化しているいくつかの課題に触れておくことにする。

学校評価の実践にとっての第1の課題は、研修の促進による学校自己評価と外部評価の強化である。欧米諸国の経験でも学校自己評価の定着に研修活動が重要な役割を果たしたことが報告されているが（Simons 2002）、国内でも広島県では研修重視の学校評価を進めてきており、それなりの成果を挙げている（石田 2006）。学校行事が多く組まれ、教職員が時間に終われる中で、多数の教員の時間を同時に研修に割くことはかならずしも容易ではないか、中間評価や最終評価のためのデータの分析を校内研修で行う等のユニークな試みも始まっている。日本の学校評価システムでは自己評価が中核であり、その質の向上なしに学校評価システムの定着はあり得ないであろう。一方、外部評価の研修についても、評価委員による評価の経験と学校に関する知識の双方の深化を図る必要があることから、委員就任の条件を3年くらいの期間とし、その貢献に対して教育委員会レベルで認知すること等が考えられて良いのではないだろうか。

第2の課題は、教育委員会の学校に対する支援である。この支援には、学校自己評価や外部評価の実施に対する指導・助言と学校による評価の結果に基づく改善の支援とが考えられる。前者は、各学校に欠けている評価の技術的な理解を補うためであり、後者は実施した評価が確実にPDCAサイクルのアクションの段階につながるためである。この支援に関する現時点での問題は、教育委員会の指導主事がそこで必要とされる役割を果たしていないことにある。各学校の評価委員会に所

属する教員は評価の経験を積んでおり、そのような経験をもたない教科担当の指導主事では指導も助言もできなくなりつつある。また、そのような関係の下では、指導主事が担当校の必要とする支援内容を教育委員会に伝えることもできない。そこで、今緊急に必要とされるのは、教育委員会の指導主事に集中的に学校評価の研修機会を提供することである。パワーアップした指導主事が学校に対して友好的直言居士の役割を果たすようになる時、学校評価が学校と教育委員会の対話のツールとして活かされるようになるであろう。

第3の課題は、学校評価活動への地域の参加である。繰り返し述べたように学校評価システム構築の目的は、学校が地域に開かれ、信頼されるようにすることである。ほとんどの学校でこの点は学校評価の中に重要評価項目として取り入れられている。しかし、その中身は「評価結果をホームページや学校便りで伝える」、「地域の人材を教育活動に登用する」、「地域住民が児童生徒の安全のためのボランティア活動を行う」、「教職員が地域の活動に参加する」等で、評価がツールとして活かされていない。学校評価は学校と学校関係者が「データに基づく対話」を深めるためのツールであるとするなら、地域関連の項目については、地域が学校自己評価計画作成に関与し、評価項目・評価指標の選択や評価基準の決定に直接参加してはどうだろうか。保護者・地域による学校行事への参加率が低下しているなら、その改善策の模索は、学校評価の枠組み作りの一環として学校と保護者・地域画共同で行えばよいであろう。

第4の課題は、学校評価に関する専門的人材の不足である。全国的規模で第三者評価の実施が考えられているにも拘わらず、日本には大学院レベルで評価教育を行っている教育機関がないのが現状である。全国の学校で多数の教員が学校評価の実践を通して組織評価、事業評価の知見を蓄積していく時、必ずその中から専門的な知識と技能に対する要請が出てくるであろうことは容易に想像できる。また、高等教育機関で、専門的な研究教育を行わなければ、そのような要請に応えることはできないであろう。大学院レベルでの評価教育・研究の開始は学校評価だけでなく大学評価さらには他の分野の評価にとっても緊急の課題とな

っている。

最後に、日本評価学会は、学校評価研修プログラムを対象とする認証制度を発足したことにより、前述した学校評価の研修機会の拡充を支援する体制を確立した。既に認証第1号を受けた広島県立教育センターに続いて、他の教育委員会や教育センターもその制度の活用に関心を示している。学会がそのような関心の拡大に応えるには、単に認証機能を発揮するだけでなく、認証取得の過程を技術的に支援することも必要である。例えば、専門的な学校評価研修講座を立ち上げる際の講師の紹介や教材の提供である。学会会員の活躍の舞台を広げることは、学会自体の活性化にも資するものであろう。

注記

- 1 数少ない例外のひとつに長尾・和佐・大脇（2003）があるが、それも学校評価の理論的基盤について概説しているに過ぎない。
- 2 例えばフランスの国民教育省では、年次的に全国レベルでの視学報告的評価調査を行う。窪田・木岡（2004）の第V章参照。
- 3 例えば、オハイオ州教育省（2006）の説明責任システムでは、小学3年生以上高校2年生までの教科学力テストの結果と卒業比率・出席率の計30指標のデータを、ホームページで各学区・各学校ごとに公開している。
- 4 米国の認証制度は、全国を6地域に分けてそれぞれ独自に組織運営する学校認証委員会とそれらが共同で設立した全国組織（National Study of School Evaluation/NSSE）から成り立っている。その仕組みと活動については、NSSE（2006）のホームページを参照されたい。
- 5 欧米諸国の学校評価制度については、窪田・木岡（2004）、Nevo（2002）に詳しい。
- 6 日本における学校評価の歴史的展開については、木岡（2003）の第2章を参照されたい。
- 7 以上の文科省の動きを中心とする展開については、文部科学省（2006）を参照されたい。
- 8 学校による自己評価や外部評価の実施事例については、広島市教育委員会（2005）、窪田等編（2005）を参照されたい。
- 9 英国の学校評価システムについては、窪田・木岡（2004）の第II章参照に詳しい。
- 10 広島県における学校評価の現状と課題については長尾（近刊）を参照されたい。
- 11 広島市教育委員会のホームページで検討内容を公開している。（<http://www.city.hiroshima.jp/icity/browser?ActionCode=genlist&GenreID=1000000001941>）
- 12 以下の評価手法の理論は、長尾・久慈・広島県立教育センター（2007）に依拠している。
- 13 ロジック・モデルについては、Rogers et.al.（2000）を参照されたい。

参考文献

- 石田謙豪（2006）「広島県における研修重視の学校評価の取り組み」、日本評価学会東北支部設立記念シンポジウムにおける実践報告（『学校評価・政策評価の課題と展望 報告資料集』15-29）
- 金子郁容（2003）『学校評価 - 方法共有のデザインとツール』、ちくま書房
- 木岡一明（2003）『新しい学校評価と組織マネジメント』、第一法規
- 木岡一明（2004）『学校評価の「問題」を読み解く - 学校の潜在力の触発』、教育出版
- 窪田眞二・加須南小学校学校評価委員会（2005）『学校第三者評価の進め方』、学陽書房
- 窪田眞二・木岡一明編著（2004）『学校評価のしくみをどう創るか - 先進5カ国に学ぶ自立性の育て方』、学陽書房
- 小島宏（2004）『学校の外部評価と説明責任 - 担任、主幹/主任、管理職必携』、明治図書
- 中留武昭（1999）『学校経営の改革戦略 - 日米の比較経営文化論』、玉川大学出版部
- 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）：57-68
- 長尾眞文（近刊）『学校評価』、三好皓一編『評価論を学ぶ人のために』、世界思想社
- 長尾眞文・久慈恵子・広島県立教育センター（2007）『学校評価連携推進講座テキスト』、広島県立教育センター専門研修講座、東広島市
- 長尾彰夫・和佐眞宏（2003）『大学と教育委員会のコラボレーションの可能性』、長尾彰夫・和佐眞宏・大脇康弘（2003）『学校評価を共に創る - 学校・教

- 委・大学のコラボレーション』、学事出版、181-196
- 西村・天笠・堀井編(2004)『新・学校評価の論理と実践 - 外部評価の活用と内部評価の充実』、教育出版
- 伯井美德(2006)「学校教育システムの構築について」『教育委員会月報』、2006年7月号、16-19
- 広島市教育委員会(2005)『学校評価実践事例集』、広島市
- 文部科学省(2006)『義務教育諸学校における学校評価ガイドライン』、東京
- 文部科学省初等中等教育局(2006)『学校評価及び情報提供の実施状況調査結果の概要(平成16年度間調査結果)』(http://www.mext.go.jp/b_menu/houdou/18/01/06011710.htm)
- Compton, D. W. , Baizerman, M., & Stockdill, S. F. (2002). *The Art, Craft, and Science of Evaluation Capacity Building*. New Directions for Evaluation, San Francisco, CA: Jossey-Bass Publishers.
- Conner, R. (2001). *Evaluation Training in the United States: A Description and Comparison of three Representative Programs* (mimeo). Paper prepared for the International Development Center of Japan, Tokyo.
- Doherty, J., MacBeath, J., Jardine, S., Smith, I., and McCall, J. (2001). *Do schools need critical friend*. In MacBeath, J. and Mortimore, P. *Improving School Effectiveness*. Buckingham, The Open Press, 138-151.
- Hopkins, D. and West, Mel. (2002). *Evaluation as school improvement: A developmental perspective from England*. In Nevo, D. (ed.), *School-based Evaluation: An International Perspective*. Oxford: Elsevier Science Ltd, 89-112.
- Joint Committee on Standards for Educational Evaluation (1994). *The Program Evaluation Standards*. Thousand Oaks, CA: Sage.
- Love, A. (1991). *Internal Evaluation: Building Organizations from Within*. Newbury Park, CA: Sage Publications.
- MacBeath, J. & McGlynn, A. (2002). *Self-evaluation: What's in it for schools?* London: Routledge Falmer.
- NSSE (2006). *National Study of School Evaluation* (<https://www.nsse.org/aboutnsse/index.cfm>)
- Nevo, D. (ed.) (2002). *School-based Evaluation: An International Perspective*. Oxford: Elsevier Science Ltd.
- Ohio Department of Education(2006) *Ohio Accountability System* (<http://www.ilrc.ode.state.oh.us>).
- 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
- Rogers, P. J., Hacsı, T. A., Petrosino, A. and Huebner, E. A. (ed.) (2000). *Program Theory in Evaluation: Challenges and Opportunities*. New Directions in Evaluation, No. 87, San Francisco: Jossey-Bass.
- Sanders, J. R., Horn, J. L., Thomas, R. A., Tuckett, D. M., and Yang, H. (1995). *A Model for School Evaluation*. CREATE. The Evaluation Center, Western Michigan University, Kalamazoo, Michigan.
- Scriven, M. (1991) *Evaluation Thesaurus*, Fourth Edition. Newbury Park, CA: Sage Publications.
- Shadish, W. R., Cook, T. D. and Leviton, L. C. (1991). *Foundations of Program Evaluation: Theories of Practice*. Newbury Park, CA: Sage Publications.
- Simons, H. (2002). School Self-Evaluation in a Democracy. In Nevo, D. (ed.), (2002) *School-based Evaluation: An International Perspective*, Oxford: Elsevier Science Ltd, 17-34.
- Stufflebeam, D. L., (2000). The methodology of metaevaluation. In Stufflebeam, D., Nadas, G., and Kellaghan, T. (eds.), *Evaluation Models: Viewpoints on Educational and Human Services Evaluation*, Second Edition. Boston: Kluwer Academic Publishers, 279-317.

(2007.2.26受理)

参考資料

学校評価：メタ評価チェックシート

学校名： _____

日時： _____

メタ評価者名： _____

<メタ評価の方法>

下記の各チェック項目について4段階評価（優・良・可・不可）する。

各評価カテゴリーごとに、下記の方法で評点を計算する。

優の数 _____ × 3 = _____

良の数 _____ × 2 = _____

可の数 _____ × 1 = _____

評点合計 _____ = _____

各カテゴリーごとに、必要に応じて評価コメントを添える。

最後にカテゴリー別評点の合計点を加えて、総合点（75点満点）を求め、総括評価コメントを添える。

1 評価の前提条件について

1.1 学校長が学校評価の実施について明確な意思表示をしていたか？ []

1.2 学校内で学校評価実施体制が確立されていたか？ []

1.3 学校ニーズが十分に把握されていたか？ []

1.4 学校経営目標が現実的に設定されていたか？ []

1.5 学校経営目標と達成手段・活動・投入が合理的に関連づけられていたか？ []

「第1カテゴリー」評点：

優の数 _____ × 3 = _____

良の数 _____ × 2 = _____

可の数 _____ × 1 = _____

評点合計 _____ = _____ /15点

コメント：

2 評価枠組について

2.1 学校経営目標が適当な成果目標と努力目標として表現されていたか？ []

2.2 具体的、客観的で、検証可能な評価指標が設定されていたか？ []

2.3 判断基準が適切に設定されていたか？ []

2.4 生徒、保護者の意見を参考にする等多面的な評価を心がけていたか？ []

2.5 評価システムの信頼性を高める工夫（外部評価の活用等）が組み込まれたか？ []

「第2カテゴリー」評点：

優の数 _____ × 3 = _____
 良の数 _____ × 2 = _____
 可の数 _____ × 1 = _____
 評点合計 = _____ /15点

コメント：

3 評価実施プロセスについて

- 3.1 評価のタイミング、期間、スケジュールは適切であったか？ []
- 3.2 中に学校経営計画の進捗状況の適切な把握（必要な修正）を行ったか？ []
- 3.3 量的、質的に十分なデータを収集したか？ []
- 3.4 学内研修等を通して、評価の実施に関する意識統一や評価方法に関する共通理解の徹底が図られたか？ []
- 3.5 学校評価の実施方法の改善のための具体的な施策を実施したか？ []

「第3カテゴリー」評点：

優の数 _____ × 3 = _____
 良の数 _____ × 2 = _____
 可の数 _____ × 1 = _____
 評点合計 = _____ /15点

コメント：

4 分析と結論・提言について

- 4.1 分析方法や結論の導き方、制約等について明確な説明をしているか？ []
- 4.2 成果目標・努力目標の達成度について明確に表現しているか？ []
- 4.3 目標の達成度に関する十分な説明（貢献・阻害要因等）がされているか？ []
- 4.4 評価結果の分析に基づいた結論・提言が導かれているか？ []
- 4.5 将来的な学校経営の改善につながる提言が提示されているか？ []

「第4カテゴリー」評点：

優の数 _____ × 3 = _____
 良の数 _____ × 2 = _____
 可の数 _____ × 1 = _____
 評点合計 = _____ /15点

コメント：

5 結果の報告・公表と活用について

- 5.1 わかりやすい報告が作成されたか？ []
- 5.2 報告は学校評価の結果を過不足なく伝えているか？ []
- 5.3 学校全体で評価結果を共有する努力がなされたか？ []
- 5.4 評価結果の公表について十分な努力がなされたか？ []
- 5.5 評価結果の積極的な活用のための具体的措置がとられたか？ []

「第5カテゴリー」評点：

優の数 ____ × 3 = ____

良の数 ____ × 2 = ____

可の数 ____ × 1 = ____

評点合計 = /15点

コメント：

総合評点：

優の数 ____ × 3 = ____

良の数 ____ × 2 = ____

可の数 ____ × 1 = ____

不可の数 ____

評点合計 = /75点

総括評価コメント（学校側の現在の学校評価の取り組みの姿勢と将来的な展開の可能性に関する総括的な観察・判断、より良い学校評価とするための改善提案等）：

メモ：メタ評価実施方法（使用資料/聞き取り調査/実施上の制約等）

（出所）筆者作成

Theoretical and Practical Issues in School Evaluation

Masafumi Nagao

Hiroshima University
nagaom@hiroshima-u.ac.jp

Abstract

Since the Central Education Council of the Ministry of Education, Culture, Sports, Science and Technology made a suggestion, in its report on local educational administration in 1998, that schools should carry out a self-evaluation annually and make public its results so as to make them more open and accountable to their respective communities, public schools all over Japan have been establishing a school evaluation system. Today self-evaluation by teachers is practiced by more than 95% of public primary and lower secondary schools in Japan. However, the knowledge base of those dealing with evaluation tasks in schools is quite limited. The literature on school evaluation is on the increase in recent years, but much of it is no more than tracing of the practice and attempts at theoretical sorting of the practices are rarely made. In the present paper, the theory of school evaluation is depicted as consisting of a two-layered structure with a theory of school evaluation mechanism (school evaluation system) and a theory of school evaluation practice (evaluation methods). A framework is presented for analysing them with a focus on practice. The concluding section of the paper sorts out the issues in moving from the mechanism to practice.

Keywords

evaluation theory, school evaluation system,
self evaluation, external evaluation, meta-evaluation

【実践・調査報告：依頼原稿】

自己評価の活用による学校改善の実践報告

石田 謙豪

平 恵津子

住元 しのぶ

長尾 眞文

尾道市立栗原北小学校
ishidakengou@aol.com

広島大学
nagaom@hiroshima-u.ac.jp

要 約

本稿は、広島県尾道市の中規模の公立小学校による学校評価の実践事例報告である。過去3年間にわたるこの事例の特徴は、校長が強いリーダーシップを発揮し、教職員を対象とする学校評価研修の充実を図りつつ、試行錯誤を重ねて、教育の質の改善と保護者・地域に対する説明責任の遂行の両面で一定の成果を挙げたことである。外部評価の充実も含めて評価結果の信頼性を担保する方法や学校の自律性を促進するためにも必要な経営面・組織面での評価の充実等の課題は残しているものの、自己評価の活用による学校改善の試みとして価値ある示唆を含む事例となっている。

キーワード

学校改善、学校評価、ロジック・モデル、説明責任、信頼される学校

1. はじめに

文部科学省（2006）が平成18年3月に発表した学校評価ガイドラインによると、学校にとって学校評価を行う目的は、一方で「自らの教育活動その他の学校運営について、目指すべき成果やそれに向けた取組について目標を設定し、その達成状況を把握・整理し、取組の適切を検証することにより、組織的・継続的に改善すること」、また他方「自己評価及び外部評価の実施とその結果の説明・公表により、保護者、地域住民から自らの教育活動その他の学校運営に対する理解と参画を得て、信頼される開かれた学校づくりを進めること」にある。広島県では、他県に先んじて平成14年度より学校自己評価とその成果の公表を義務化して

おり、各学校がそれぞれに上記の目的の実現に向けて努力している。本稿は、同県尾道市にある中規模公立小学校における学校評価の実践報告である。この学校の実践する学校評価の特徴は、(i) 校長が学校評価システムの構築を主導していること、(ii) 校内的な学校改善の努力に関する地域・保護者の理解を得ることをシステム構築の主眼としていること、(iii) 自己評価の実施にロジック・モデルを活用していること、(iv) 全校的な研修重視の取組を実現していることの4点で、平成18年10月に文部科学省初等中等教育局が試行した第三者評価に志願して対象校となり、調査団から校長を中心として学校評価を生かした円滑な学校経営を実現しているとの高い評価を獲得している。以下の報告では、学校評価の取組の特徴的

な点について概説するとともに、可能な範囲での取組で得られた成果について言及する。共同で報告の執筆に当たったのは、この学校の校長と教務部主任、研究主任を務める教員及び校長の依頼で評価システム構築に助言している大学教官の4名である。

1. 学校の状況と自己評価の取組の展開

(1) 学校の状況

尾道市立栗原北小学校は、広島県尾道市のほぼ中央、JR新尾道駅の北側に位置している。平成18年5月1日現在で、学級数14、児童数365名、職員数23名の中規模校である。この地域は元々農業を主産業とするのどかな田園地帯を形成していたが、南北に走る国道が東西に掛かる福山市へのバイパスや山陽自動車道と接続したことから急速に住宅地としての開発が進んだ。栗原北小学校は、宅地化に伴う児童数増により、昭和55年に近隣の小学校から分離独立し開校した。住民の多くは全国各地からの転入者であり、地域の住民同士つながりは比較的弱い。住居は一戸建てもあるが家賃の安価な民間アパートが多数あり、家庭数269の2割の家庭の児童が市から就学援助を受けている。また、親の離婚率が高く20代の親が多いという特徴がある。校区には養護施設があり、毎年30人前後の児童が施設から通学している。

学校教育に対する親の期待は基本的なしつけや保護者の個人的な都合によるものが多い。児童の課題は基本的な生活習慣が出来ていないことと学力が低いことである。生活指導上の課題が多く、学校規模に比して男性職員の割合が多い。職員の平均年齢は37歳である。本稿の共同執筆者でもある校長は、平成16年4月に県内の他都市から転任してきたが、着任時に前任校長及び教頭から課題として引き継いだのは、「本校は地域・保護者からの苦情が非常に多く、その対応に苦慮する」「学校の運営に保護者の協力を得ることが難しい」「児童の問題行動が多い」の3点であった。

(2) 自己評価の取組の展開

平成16年に赴任した現校長は、学校経営会議（校長、教頭、教務主任・研究主任・生徒指導主事・保健主事の6名で構成）¹で検討した内容と教育活動が連鎖していないこと、また教職員が個人としての動きに比して一体化した取組に消極的なことを問題視し、その改善のため、分掌単位の責任体制の必要、教職員の経営参画意識の必要、計画性導入の必要の3つを学校経営の主要課題として提起。自己評価を活用して学校教育・経営の改善を図る方針を打ち出した²。

前任校長の下での平成15年度までの学校評価は、校長が作成した学校評価表（以下、評価表と略す）に基づき、年度末に教職員が評価した結果を地域や保護者等に公表するというものであったが、経営の改善に用いられることは少なかった。教職員の学校評価に関する認識から改めるため、校長は主任をはじめ教職員に対して、学校評価の目的は、学校の現状を見直して教育活動の改善を図ることにあり、評価の結果がどうであるかよりも、結果を受けて教職員が「何を」、「いつまでに」、「どうするのか」を明確にし、進捗状況等を保護者や地域に伝える必要があること、それが説明責任を果たすことに相当すること、そしてそのプロセスに学校を挙げて取り組む必要があることを繰り返し説いた。

そこで平成16年度からは、教育活動の実践計画を評価表と連鎖させるとともに、学期ごとに実施プロセスの評価を行い、進捗を保護者や地域に発信するようにした。これまでも各学期末には各分掌・学年・学級の作成した資料に基づく教育活動の反省を行っていたが、それを学校評価集計表のみに絞って行うように変更した。その理由は、目標を意識したまとめをすることにより、見直しの内容を明確にするためである。集計表の基礎データを得るために、評価表を基に作成し学期末に実施した児童アンケートと教職員アンケートの結果を用いた。

平成17年度の学校評価については、計画作成の段階で校長が全体案を提示し、それに基づいて各評価項目について担当教員が修正を加える形で自己評価計画を作成するとともに、評価の実施につ

いては、学期ごとに各教員が評価の根拠となるデータを収集するようにした。全校的評価の原案を作るのは教務主任の仕事で、各教員から提出された結果を集計して全体的な達成度を明らかにし、次の学期に向けた改善案をまとめた。尾道市の公立小中学校の場合、評価項目の大枠は、同市教育委員会から与えられており、評価計画の詳細は学校がどれだけ丁寧な対応を図るかに依る。栗原北小は、学校評価を意識的に活用した学校経営を心がけた極めてユニークな事例と言えるであろう。

学校評価結果の公表については、取組を従前と変更したのは、説明の時期と内容である。時期は地域・保護者から求められる前に発信すること、内容は既に発信している事柄についての進捗を説明すること、また校長の説明と学年の発信内容に一貫性を持たせることである。変更した理由は学校が真摯に活動し、その状況を地域・保護者に定期的に発信する姿勢を示すためである。説明時の留意点は2つある。一つは、学校が「やろうとしていること」「現在、行っていること」の意義・目的を伝えるとともに、それがどのような状況にあるのかという進捗についての説明をすることで、進捗についての説明はそれまでほとんどなされていなかった。もう一つは、保護者から求められる内容を実行することができない場合に、根拠を示してその理由を説明しきることである。例を挙げると、校長が就任した年に保護者から出された「運動会に国旗を掲揚することに反対する」意見に対して、「学校では、オリンピックや国際会議等、国と国とが対等の立場で交流する際の国際マナーの一つである国旗の取り扱い（自分の国や他の国の国旗を敬い、尊重する態度を身につけさせるために掲揚するとともに、そのマナー）についての指導をしています。これは学習指導要領に基づくものです。運動会はそのための指導の機会の一つと考えております」³と、正面から向き合って対応することである。

広島県教育委員会（2002）の学校評価ガイドラインでは、学校は自己評価の信頼性と透明性を確保するため外部評価を実施することとしている。栗原北小では、平成16年度から3名の学校評議員に自己評価結果を提示して、意見・助言の依頼を行ってきたが、平成18年度からは文部科学省の学

校評価ガイドラインに基づく尾道市教育委員会の指導により、3名の外部評価委員が年間計画に基づく外部評価を行うように成っている。

栗原北小の学校経営は、校長がリーダーシップを発揮して、学校評価を意識的に活用することにより大きく変化し、当初校長の下に毎月20件程寄せられていた「教育活動のねらいについて」、「教育活動から生じた児童間のトラブル」、「教職員の言動」等に関する保護者・地域からの苦情が、校長の就任2年目には殆どなくなった。その間、学校評価の実施方法も進化を遂げたが、その基本的特徴は、ロジック・モデルを活用した自己評価にある。

3. ロジック・モデルに基づく自己評価システム

(1) ロジック・モデルの導入

後述するように、栗原北小の校長は、就任1年目の夏に広島県立教育センターが主催する学校評価（4日間）研修講座を受講し、ロジック・モデルに基づく学校評価について研修した。帰校後の校内研修でロジック・モデルの活用について教職員による理解の共有を図るとともに、平成17年度、18年度にもそれぞれ1名の教員を同研修講座に派遣し、それを実践に生かす条件を整えた。

ロジック・モデルは、学校が教育活動を通して達成しようとする目標とその実現のための取組の間に想定する因果関係を明示化し、適当な指標を用いてその達成度を評価するためのツールである（長尾 2007）。校長は、教職員が学校評価が何のために必要なかを十分に理解できない理由を評価の理論的理解の欠如にあると考えていた。ロジック・モデルの概念構成を使うことにより、学校経営計画に対する自己評価の枠組みを作成し、プロセス評価や成果評価を客観的かつ厳密に行うことができる。校長は、自らロジックモデルによる学校評価の枠組みを作成して、学校経営会議に提示し、自己評価表との関係を説明した。多くの教員は、自らの経験から教育活動について「こうすれば、こうなる」という想定のもとに取り組んで

いるが、それだけでは目標と活動の連鎖が弱く、評価のためのデータが精確でなくなることが多い。校長は、ロジック・モデルの活用により、そのあいまいさを克服できることを説明し、教職員の理解を得た。

ロジックモデルの活用は、目標と活動との連鎖の明示化を促すとともに、経営の問題点を明らかにすることも可能にした。それにより学校が進むべき方向について、教職員間の意見の一致を図ることができるようになった。また、管理職が進めようとしている学校評価の枠組みを教職員が理解するとともに、全員の意識を同方向に向けるためにも有効であった。そして、経営の質の改善や課題に対して意図的・計画的に取り組むことも可能になった。導入の過程では、途中で定期的かつ多面的な評価を心がけ、それを支える客観的なデータの収集に特別留意する必要があった。後述するようにかなり校内研修や外部からの技術的支援に頼ったのも事実である。しかし、教職員全員でロジック・モデルを使った自己評価について習熟する努力を重ねた結果、現在ではかなり明確な理論的検討に裏打ちされた自己評価を実施できるにいたった。

表1は栗原北小の平成18年度の学校経営計画、表2はそれに基づく学校評価総括表（11月現在）を示したものである。年間目標に相当する短期経営目標は「確かな学力の向上」が4項目、「豊かな心の育成」が2項目、「健やかな体の育成」が2項目、「秩序づくり」が3項目と計11項目に絞られている。ロジック・モデルによる検討を経て、それぞれの目標についてその達成のための方策が特定されている。さらに、各目標について、成果の達成度を評価するための指標が指定され、7月、11月、1月の成果目標値が設定される。例えば、確かな学力の向上を図るための基礎学力の習得については、全校的に週5日ドリルに取り組ませることを目標として掲げ、取組としては、計画的実施、繰り返し学習、ミニテストの実施、生徒ごとの目標設定が想定されている。これらの取組の成果の判定は、各チェック時期にミニテストで80点以上の児童の割合が80%に達することを目安に設定する。表によると、平成18年度の11月チェックでは全校的な平均点が78点で、7月チェックより上昇している

が、まだ目標を僅かに下回っている。年度末には、年間達成度を計算し、目標値に達したならA評価として、次期には新たな目標を設定し、それに満たない場合には、達成度に応じて取組の実施を継続するなり、目標自体の再考をする。

(2) 学校経営の進行管理と評価データの分析

栗原北小では、学校経営の進行管理を自己評価の重点目標の達成度を継続的に追跡することで行っている。この目標には、テストの点が80点以上の児童の割合や呼び捨てにされていない児童の割合といった成果目標もドリルを週5日したかや授業で「さん」、「くん」と呼んでいるかといった取組目標も含まれ、年初に決められた重点目標の達成シートが作られる。追跡の主要な手段は、教員と生徒による毎月末の「振り返り」で、教員と生徒がそれぞれ過去1ヶ月それぞれの目標をどれくらい達成できたかをパーセントで表現する。教員の場合は、担任教師の自らの担当学級に関する振り返りであり、児童については担任教員が挙手等の方法で学級の達成度のデータを出す。教務主任は、全データをまとめ、特に教員と生徒の振り返りデータを付け合せることにより、計画が予定通りに進行しているかを判断することができる。この簡便なデータ創出方法は、アンケートのように手間が掛からず、しかも継続的にデータが得られるので、極めて有効である。

表3は、そのような振り返りデータの分析例である。上段は、「週5日ドリルに真剣に取り組ませる」目標について、11月時点での進行管理を行ったものである。そこでは、前述したように、成果目標がテストで80点以上取らせることに関する教員の振り返りデータで表現（78%）されているが、取組目標についても教師のドリル学習を週5日実施するに関する判断と児童による朝のドリルを一生懸命しましたかに関する判断とが並べて提示されており、6月時点でのデータと比較して結果を考察し、3学期における対応を述べている。成果目標と取組目標の組み合わせと教員と生徒・児童の評価の連動は、自己評価の信頼性を高める上で極めて重要である。参考までに、下段に友達を呼び捨てにさせないと言う目標に関する振り返りデ

ータと考察を示しておく。

教育活動の進行管理にロジック・モデルを活用する試みは実際に多様な形で展開しており、それ自体ロジック・モデルのツールとしての定着を示唆している。例えば、表5は、研究主任が研究活動推進の運営のために作成したロジック・モデルで、これを用いて進行管理を行っている。

(3) 評価結果の公表

栗原北小の自己評価システムの導入は、学校経営の取組について保護者・地域のより良い理解を取り付けたいという校長の強い意向を出発点としている。前にも述べたように、そのための学校側の方策は、学校に関する情報提供を出来るだけ前倒しで行うことである。学校自己評価の結果についても、年度末を待つことなく、一年を通して随時実施する。学校は、下記の通り複数の情報伝達手段を持っている。

校長室便り：

校長の経営方針、「何のために」「何を」「いつまでに」「どうしようとしているのか」その「進捗」状況を月に一度、校長から発信するもの。配布先は保護者・地域の公的施設・学校評議員・民生児童委員・町内会役員・隣接学校である。

栗北通信：

校長室便りを受け教頭が毎月、中旬までに発行するものである。内容は校長室便りの内容の進捗の具体化であり、学校としてのとらえが主。配布先は校長室便りと同様。

学年便り：

校長室便りを受け、学年が月初めと半ばに発行するもの。内容は、学年としての月ごとの取組と、その結果が主なもの。配布先は学年の各家庭。

このような複数の情報伝達手段を使って継続的に学校の取組について説明してきた結果、学校側としてはある程度の効果が得られたと考えている。そう考える理由は、学校の発信する説明に対して肯定的な見方が出てきたことである。これは、

校長の発信と学年の発信内容に一貫性を持たせたことで学校が組織として活動していることを伝えることが出来つつあるからである。また、既に発信している内容についての進捗を説明してきたことで学校は確かに活動しているということが伝わっているからである。

厳密な根拠データではないが、保護者・地域の委員、来校業者等に学期に1回、行事（運動会・公開研究会・安全マップ作り）を行った時に実施しているアンケートがそのことを示唆している。アンケートの対象者は校長室便りの配布先と同じである。その内容は多肢選択と記述を併用したもので、来校業者等にはその場で記述してもらい、保護者と地域の委員等の回収は後日、児童を通じて行っている。回収率は通常地域の委員は100%、保護者はおよそ70%である。それによると、下記のような反応が得られている。

経営のビジョンが明確である。(10名)

学校がやろうとしていること、していることが良く分かるようになった。(10名)

学校が何に取り組み、その課題をどのようにとらえているか、家庭では何が必要なのかが分かり、学校との連携の大切さを痛感している。(5名)

家で子どもと学校のことをよく話すようになった。(3名)

当初のアンケートでは、学校に対する非難の記述がほとんどであったが、18年度にはほとんどなくなった。

4. 研修重視の取組

栗原北小の学校評価の取組に最も特徴的な点は、校長が率先して自ら学校評価について学習するなど、研修を重視していることである。主要な研修及び関連活動を挙げるだけでも、下記が挙げられる。

平成16年度：

校長が広島県立教育センター主催の「学校評価(4日間)研修講座」受講

大学教官による全教職員対象の学校評価講義

平成17年度：

研究主任が広島県立教育センター主催の「学校評価（4日間）研修講座」受講

日本評価学会派遣の「評価インターン（院生）」受け入れ・研修実施

平成18年度：

中堅教員が広島県立教育センター主催の「学校評価（4日間）研修講座」受講

評価インターンとの継続的連携文部科学省第三者評価制度試行調査の受け入れ

ここで挙げているのは全て、外部の研修プログラムに参加するか外部の専門家人材を活用して研修を行うものであるが、それ以外に自校でも校内研修を多数回開催している。校長によると、そのような研修からは、(i) 学校経営の問題点を明らかにする、(ii) 教職員間の思考ベクトルの一致を助長する、(iii) 経営参画意識を向上させる、(iv) 学校評価手法の理解を深めさせる、といった効果が期待される。(iv)を除くといずれも一般的な観点から見た効果の発現で、必ずしも研修の効果とは言えない。

(iv)の評価手法の理解について、例えばロジック・モデルについて教職員はどれくらい理解しているのだろうか。この点については、平成17年12月に栗原北小で全教職員を対象としてロジック・モデルに関する研修を実施した日本評価学会派遣の評価インターンが興味深いデータを残している。研修終了後の研修に関する評価アンケート(串本・渡辺 2006)で、「ロジック・モデルはある行為に関する原因と結果の関係を簡略的に明示化したものです。ロジック・モデルの資源(投入) - 活動 - 産出 - 成果の流れを理解できましたか」と問うたところ、「強くそう思う4点 - 全くそう思わない1点」のスケール回答で、2.79と相対的に低い理解度を示した⁴。しかし、その研修を通じて、一人を除いて全教員がロジック・モデルを理解できたとし、さらにロジック・モデルを自らの教育現場で活かすことが出来ると回答していることから、研修は組み方によって確実に学校評価の実践にプラスの効果を持ち得るものと思われる。

5. 結びに代えて

栗原北小は、校長の強いリーダーシップと全教職員が研修活動に参加する懐の深い取組により、比較的短期間に学校自己評価システムを定着させた。その方法的な特徴は、ロジック・モデルを活用して学校が行う教育・経営活動の目標と取組の連鎖を因果関係軸で明示的に捉えようとしているところにある。方法の理解に未だ不十分なところがあるにせよ、それを軸として学校経営計画の学校評価表への落とし込みや一年を通しての学校経営・教育活動の進行管理を行う慣行は十分確立されているように思える。

学校評価の活用を通して学校経営の改善を図っていく上で、栗原北小の経験が示唆する教訓の第一は、教職員の経営参画意識の向上であろう。学校経営の目標を受けて各分掌の主任が分掌運営計画のロジック・モデルを作成し(表4は研究主任の作業結果の例示)、その進行管理を行うことにより全ての教育活動と経営目標が整合的に生まれ、教職員の一体感が醸成される。それに関連して、校長が教職員の経営参画意識の向上の鍵を握るとされる主任クラスを対象に、5月と12月にアンケートをして、学校評価の目的とロジック・モデル導入の目的に関する理解度を聞いたところ、5月には50%強であったものが、12月にはいずれも80%を超えた。同様に組織的な学校経営の取組の必要性(例えば、学校の重点的な取組の進捗状況の発信や学校経営会議の役割)の理解についても、明らかに主任クラスの意識の変容が認められた。さらに、学校経営会議における主任クラスの発言でも、下記の点が確かめられた。

各分掌の主任が年間の経営計画を作成しそれに基づいた進行管理を行うので検証が容易であり、計画的な仕事ができる。(7/7名)

月ごとの重点を決めて取り組むので指導が徹底する。(7/7名)

検証データが客観的なので説得力がある。

(5/7名)

栗原北小が学校評価の取組を本格的に開始した平成16年度の経営的な問題が、保護者・地域との

関係構築であったことを考えると、同校の学校評価経験は極めて示唆に富むと言える。校長を先頭に、学校は自らの取組を保護者・地域に積極的に説明することで経営に対する理解を得、あわせて経営の改善を図る試みについて協力を求めた。学校の取組に対する理解を得るために重要なことは学校としての一貫した説明を行うことである。栗原北小の経験は、厳しい条件下に置かれている学校でも、児童をこのようにしたいというビジョンを示した後に適切に教育活動の進捗状況の説明を重ねれば、保護者・地域からの積極的な反応が得られることの一例である。校長は、児童に対する日常的な問いかけの中で、「学校からのお便りを必ず持って帰って家の人に見せる」、「家で学校のことをよく話すようになった」と答える児童が増えていることから、このような学校状況の好転を確信している。

以上、栗原北小の自己評価を中心とする学校評価の実践事例について報告したが、最後に同校が抱えている課題を挙げておきたい。一つは、学校評価システムの定着がもたらす経営と教育の質の改善効果の客観的な説明及び検証である。それは学校教育の継続的改善の証左ともなるが、同時に学校が果たすべき説明責任の遂行の中核を成すべきものである。また、学校評価が学校と地域・保護者・教育委員会等の学校関係者とのデータに基づく対話を促進するツールとして機能することが期待されるなら、この関係のあり方に関するデータを創出する手段として地域・保護者・教職員を対象に実施しているアンケートの内容・回収方法等の充実を図る必要がある。

もう一つの課題は、現在の学校評価の取組が教育活動に偏していることで、教育の地方分権化の促進で学校の裁量権が拡大していくことが想定さ

れているなら、学校経営の組織的観点により大きな比重を置く評価のあり方も検討すべきであろう。経営という視点では、人・金・物も含まれる。予算の執行状況、とりわけ旅費の用途については校長のビジョン・目標に対してどのような経営が行われているかを測るバロメータとなる。この点については、栗原北小の現実の課題である外部評価体制の充実とあわせて県・市教育委員会あるいは国が検討している第三者評価制度のあり方とそれに拘わる学校支援体制の構築にも大きな期待がかかる。

注記

- 1 現在は特別支援コーディネーターが加わり7名で構成。
- 2 ここでの議論の詳細については石田（2006）を参照されたい。
- 3 平成16年6月9日付「校長室便り」から引用。
- 4 評価目的の多様性や測定と評価の違いに対する理解等は、同じスケールで3.32、3.26である。

参考文献

- 石田謙豪（2006）「小学校における校長主導の学校評価の事例」『日本評価学会春季第3回全国大会報告集録』
- 長尾眞文（2007）「学校評価の理論と実践の課題」『日本評価研究』
- 広島県教育委員会（2002）『広島県における学校評価システムの在り方』、広島県学校評価システム検討会議報告書、広島市
- 文部科学省（2006）『義務教育諸学校における学校評価ガイドライン』、東京都

（2007.2.26受理）

表1 平成18年度 尾道市立栗原北小学校 学校経営方針

尾道市教育の方向性

夢と志を抱く子どもの育成 - 未来を抱く人間力の育成 -

学校教育目標

児童一人一人の知・徳・体の伸長を図る全人教育の創造

経営理念

出来るか出来ないかではなく、子ども達のためにするかしないか
- まず動く、動きながら考える -

経営目標

見通し・一貫性・連鎖・客観性
- 学校組織の機能化、カリキュラムの充実と説明責任 - (研究主題)

教育活動計画 (特色ある学校)

知 (確かな学力)
 目指す児童像、「自ら学び、考え、判断し、自信を持って行動できる児童」

1 学力の向上

- ・ (朝読書、習熟度別少人数指導、年間指導計画と評価評定の充実、言語技術教育・帯のドリル・早口言葉)
- ・ 学力調査に基づく学年目標の設定
- ・ 家庭学習の目標設定
- ・ 補足的・関連的学習の充実

2 県立教育センター研修等の受講と定期的な授業研究

3 機能的な小中連携

徳 (豊かな心)
 目指す児童像・「友達を思いやり、共に感動しあう豊かな心を持つ児童」

- 1 積極的に関わる生徒指導の充実
- 2 学年・学級経営案の作成と実践・評価
- 3 学校評価表と連鎖した児童会活動
- 4 道徳の授業研究
- 5 感動のある教育活動の創造

体 (たくましい体力と健康)
 目指す児童像、「たくましく生きるための体力や健康づくりに努力する児童」

- 1 体力テストの実施と分析・改善・指導
- 2 体育の指導内容の見なおし
- 3 食指導の充実 (委員会活動との連鎖)

経営計画 (信頼される学校)

<p>1 説明責任</p> <ul style="list-style-type: none"> 進捗の定期的発信 説明責任を果たす説明会の実施 参観率の増加 服務規律の厳正確保 	<p>2 経営の機能化</p> <ul style="list-style-type: none"> ロジックに基づく進行管理 学年主任会議の充実 組織的な協働行動の促進 報告・連絡・相談の徹底 一目標の実践
---	---

めざす教師像
 児童のことを心配する先生・児童の長所を見つけ褒める先生・児童の不正を厳しく戒める先生・児童の自立を支援する先生

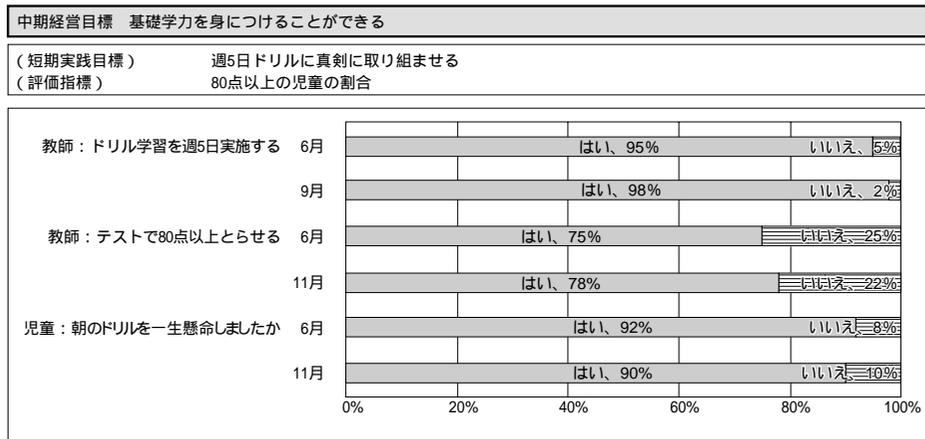
(出所) 平成18年度栗原北小学校経営構想から引用

表2 平成18年度 学校評価総括表

尾道市立栗原北小学校

経営理念	～出来るか出来ないかではなく、子どもたちのためにするかしないか～ ロジックモデルにより、各担当が進行管理を行う。									
	中期経営目標	短期経営目標	目標達成のための方策	評価指標	7月 目標値 達成値	11月 目標値 達成値	1月 目標値 達成値	達成度	評価	次年度改善策
確かな学力の向上	基礎学力を身につけることが出来る。	週6日モデルに真剣に取り組ませる。	計画的にやりきる。(授業で提出する)のテストと課題とは異なり、既習学習をくり返し定着させる。はげが学習で、漢字計算のミニテストを実施する。個々のあてを決定し、達成感を持たせる。	80以上の児童の割合	80	80	80	78	B	
		学力検査の課題に学年統一の問題で、取り組ませ、3ポイントアップさせる。	学習の約束をステップを踏んで定着させる。(1年)句読点の書き方(2年)句読点の打ち方や会話文の書き方を指導する。(3年)類似問題に継続して取り組ませる。(4年以上)	学力テストの結果3ポイントアップ	75	78	80	1月実施	1月実施	
豊かな心の育成	自らの考えを発表することができる。	自らの考えを発表させる。	学習規律を身につけさせる。自分の考えを持たせる時間を保障する。考えたことを発表する時間を保障する。	学力テストの結果3ポイントアップ	50	70	80	95	A	
		言語技術を身に付けさせる。	言語技術を計画に沿って実践する。言語技術を使った(結論先行・根拠を挙げるなど)表現方法を身に付けさせる。言語技術を教科・特活の場で活用する。学習規律を身につけさせる。自分の考えを持たせる時間を保障する。考えたことを発表する時間を保障する。	学年で計画した実施回数を充足した学級の割合 学年の目標に到達した児童の割合	66	76	80	103	A	
豊かな体の育成	思いやりのある言動ができる。	思いやりのある行動を発表させる。	正しい言葉遣いをする大切さを指導し、友達の名前を正しく呼ばせる。授業では、敬称をつけて呼ばせる。「ーさん、ーくん」と呼んでいるかを振り返らせる。	学級で決めた回数以上に報告機会を設けた学級の割合	80	80	80	90	A	
		毎日外遊びを30分以上させる。	遊びを紹介する。外遊びの場を設定する。振り返りカード	呼び捨てにされていない児童の割合	70	90	80	55	B	
秩序づくり	生活習慣を見直し健康な体を作ることが出来る。	準備しないのでバランスよく食べさせる。	食べられる量を把握し、調節する。ランチルームでの給食指導を行う。振り返りカード	毎日外遊びを30分以上した児童の割合	80	80	80	93	A	
		時間を守らせる。	教師が率先して時間を守り、規範を示す。	残食00日の割合	60	70	80	114(60)	A(B)	
秩序づくり	マナーとルールを身につける	清掃ができる。	教師が率先して清掃を行う。教師が一層に清掃を行う。	音楽を聴いて教室へ歸る児童の割合	50	70	80	113	A	
		あいさつをさせる。	分担当を細かく決め、手順を示す。担当場所を1ヶ月は変えないで、徹底させる。教師が率先して挨拶する。授業の始まりや終わりの挨拶、教室等への出入り時の挨拶など、生活の様々な場面での挨拶を指導する。	担当の仕事責任を持って仕上げる児童の割合	92	90	80	76	B	
【評価】A：8割以上 目標達成とみなし新たな目標設定 B：8割未満5割以上 8割を超えるまで継続実施 C：5割未満 目標の見直し 重点課題には、をつける (出所)平成18年度11月時点栗原北小学校学校評価総括表から引用										

表3 平成18年度学校評価 振り返りデータ分析例（11月）



結果

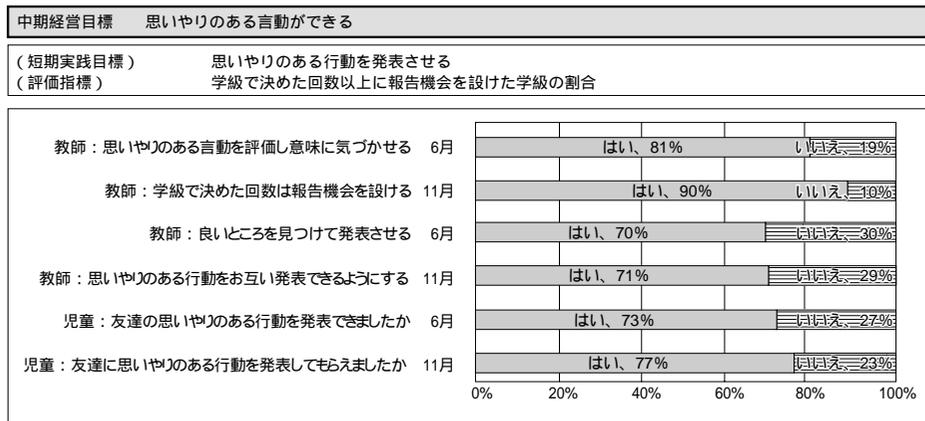
ドリル学習を週に5日行うというリズムが、9月に98%となり定着した。
 テストで80点以上とれる児童の割合が、3%増加して78%になった。
 しかし、まだ目標の80%には満たない。
 「一生懸命しましたか」という問いに、6月より2%減少したものの、90%の児童が、「はい」と答え、真剣に取り組んでいた。

考察

継続して指導することにより、少しずつではあるが着実に成果が上がっている。
 一生懸命にしない10%の児童にどう取り組むか... (要検討)

3学期に向けて

引き続き週に1回はテストを行い、学力がついているかを確認する。
 テスト結果を継続して記録させることで伸びを実感させたり、評価したりする取組が必要である。
 間違い直しを確実にさせたり、間違いの多かった問題に繰り返し取り組ませることで、定着を図る。



結果

実施に無理がないよう見直しを図った結果、学級で決めた回数以上に報告機会を設けた学級が90%になった。
 教師が思いやりのある行動を発表できるようにした(71%)結果、発表できた児童が73%、発表してもらえた児童が77%になった。目標の80%まで、あと3%である。

考察

報告機会が定着することにより、さまざまな思いやりのある行動が紹介されるようになった。

3学期に向けて

児童の肯定的な評価に努め、思いやりのある行動を教師自身が認める。
 思いやりのある行動を全校集会などで紹介することにより、児童の意識の高揚を図る。
 エンカウンターを取り入れながら、集団づくりをする。

(出所)平成18年度11月時点栗原北小学校学校評価結果の分析から引用

表4 尾道市立栗原北小学校 平成18年度研究主任ロジックモデル

目標	経営理念 出来るか出来ないかではなく、子どもたちのためにするしかないか ビジョン PDCAサイクルにそった学校運営 (事業目標) 理論に裏打ちされた実践力と組織で動く教職員集団	研究理論の共通認識	授業力をあげる (構成力、展開力、改善力)	学級経営の充実
課題	研究体制の確立 校内環境の充実	研究理論の構築	指導案作成形式、手順の提案	学級経営研修(3回)
活動	言語技術年間計画、系統表 の見直し 英語活動学年別年間計画の 見直し 言語コーナーの活用 英語コーナーの活用	研究理論の構築 理論研修12回(英5回、言5回、道2回) 研修のまとめの作成(生員レポート提出) 毎週月曜日の研修 職員委員会での言語技術演習 問答ゲーム事例集作成(夏、冬) 研究のまとめ作成	指導案作成形式、手順の提案 校内授業研究(一、二回) 校内全体研に関わる事前研(1人1回) ブロック研の検証(分析と考察) 公開研究会での検証 公開研究会の事前研修 公開研の検証(分析と考察)	望ましい人間関係づくりガイドライン作成 「栗北小学校学習規律」の定着 言語環境を整える すべての授業において、「目標の提示とまとめ」 を行う
投入	見直しの時間 見直しの時間 掲示ベースの確保、掲示物 作成時間 掲示物作成時間 掲示ベースの確保、掲示物 作成時間	まとめのために要する時間 講師招聘手続き、予算配分、研修時間 まとめに要する時間(15分×12回×人数) 研修時間の確保、集まる時間 各人の演習準備に要する時間 まとめのために要する時間 作成するために要する時間	作成時間 授業研究に係る時間 事前研修に要する時間 検証の時間 公開研究会開催 指導案作成、事前研修に要する時間 検証の時間	研修時間、講師 外部専門家による指導 指導時間、各月の振り返り、集計 配布物、掲示物、教材等の検討時間 教材研究に要する時間
想定結果	19年度言語技術年間計画、系統表の完成 19年度英語活動学年別年間計画の完成 言語コーナー貼替3回 英語コーナー貼替1回	研究理論の構築(英語、言語)(5月) 1回1時間の理論研修×12回 15分×6回、「研修のまとめ」の配布6回 英語10分×16回、言語10分×16回 一人4回ずつの演習 問答ゲーム事例集の完成 研究のまとめ完成(2月)	作成形式、手順にそった指導案の作成 授業研実施(全体14回、ブロック14回) 事前研修の実施(一人1本) 公開研究会冊子への掲載 公開研実施の成果と課題 公開研指導案集完成 研究のまとめへの掲載	研修時間2時間×3回 児童が集団の中で安心して話ができるような 教師と児童、児童相互の好ましい人間関係づ くりのガイドライン完成 各月の振り返り、集計実施 適切な言語環境の整備 指導と評価の一体化した授業
短期成果	(1)計画遂行能力の向上() (2)適切で効果的な掲示物作成能力の向上()	(1)理論に基づいた教育実践力の向上() (2)問答ゲーム問題集が活用されている() (3)研究理論の共通認識()	(1)指導案作成形式、手順の確立() (2)授業構成力、展開力、改善力の向上() () () () () ()	(1)学級経営能力の向上() (2)学習規律の定着() (3)教師の言語技術力の向上() (4)授業力の向上()
中期成果	(1)児童の実感を踏まえて、年間計画や指導案を作成し、指導のねらいを達成する (2)校内環境の美化に向け一人一人が創意工夫する	(1)理論に基づいた教育実践力の向上・定着 (2)研究理論が共通認識され、テーマに沿った議論が行われる研修会が行われる	(1)確かな授業力を身につけている (2)研修の機会等を積極的に活用し、新たに必要とされる知識や技能の習得に取り組み (3)自らの実践や研修成果を積極的に提案し、改善に努める	(1)学校経営目標や児童の実感を踏まえ、学級経営ができる。 (2)相手や場面に応じた正しいことば遣いで、指示や説明ができる。 (3)新たな課題や困難な課題に意欲的に取り組み、解決を図ろうとする。
外部条件				

(出所) 平成18年度栗原北小学校学校研究推進のロジックモデルの分析から引用

A Report on School Improvement Practice through Self-evaluation

Kengou Ishida Etsuko Taira Shinobu Sumimoto Masafumi Nagao
Graduate School of Policy and Management, Doshisha University Hiroshima University
Onomichi Municipal Kuriharakita Elementary school nagaom@hiroshima-u.ac.jp
ishidakengou@aol.com

Abstract

It is important that information that the principal gives and information that other teachers give relate, and they act as an organization, To attempt management improvement of the school by fostering the accountability.

There are two important things for accomplishing accountability.

One is to explain the situation of the progress management.

Another, when outside people request execution to the school, it is important that we finish explaining showing them grounds that cannot be done in case of difficulty.

Accomplishing accountability becomes one of the means for the school to attempt reexamination of its own management.

Moreover, for people and guardians in the region, it becomes one of the measures that they understand school management by sharing information and problems among them and the school personnel.

Keywords

school improvement, school education, logic model, accountability, reliable school

【研究論文：依頼原稿】

大学評価・学位授与機構による試行的大学評価事業の評価

齋藤 貴浩 林 隆之
 大学評価・学位授与機構 評価研究部
 tksaito@niad.ac.jp

要 約

大学評価・学位授与機構（以下「機構」と記す。）は、大学評価を行う第三者評価機関であり、2000年度から2003年度まで国公立大学及び大学共同利用機関を対象に、延べ550組織に及ぶ試行的評価を実施した。評価とは、政策、施策、事業等を体系的かつ客観的に分析し、価値判断を行うことであるが、しかし、その評価自体も、各種資源を投入して評価対象並びに社会に対して何らかの変化を生じさせる事業である。つまり、評価自体も一つの事業として評価の対象となりうる。本研究では、機構が行った試行的大学評価について、ロジックダイアグラムを基に評価事業の設計と評価結果の妥当性並びにその成果について論じ、また費用効果の観点から評価事業の効率性を検討する。さらに、新たな大学評価システムの構築に活かすため、試行的評価経験と、評価の評価によって新たに得られた教訓から提言を行う。

キーワード

大学評価、メタ評価、ロジックダイアグラム、費用効果、自己評価

1. はじめに

ここ数年というごく短い期間で、初等、中等、高等教育の別を問わず、教育への第三者評価の導入が急速に進んでいる。教育は誰もが体験し、また身近な存在であるがために、教育の評価は容易であると思われがちである。しかし、実際にはこれほど評価が困難な事業はない。例えば、教育の成果は多様であり、目に見えず、個人に依存し、広く波及し、長期的な展望を必要とし、そして極めて広範な利害関係者が相反する利害を有しつつ強く主張もしている。さらに、大学を評価すると

なれば、学術・研究という高度な専門性とその特殊性を考慮に入れる必要がある。このような特殊な事業評価を、どうすれば誰もが納得するような形で遂行できるのかという命題は、最適解があるのかどうかすらわからない。

教育に限らず、これまでに評価関係者は評価を通じてどのように政策や事業をより良くしていくかということに努力してきた。しかし、評価という事業自体もまた、人的資源や資金を必要とし、事業や組織、さらには社会を良くしていくための一つの事業に他ならない。事業である以上は、改善や説明責任の遂行を目的とする評価という活

動、それ自体にも、改善の必要や説明責任が生じている。実際に、大学の第三者評価を目的に創設された大学評価・学位授与機構（以下「機構」と記す。）による試行的大学評価では、評価が大学に大きな負担と影響とを与えることから、評価対象である大学側から常に「メタ評価（評価の評価）」の実施を求められてきた（例えば国立大学協会2002, 2003a, 2003b）。

「評価の評価」を意味するメタ評価は決して新しい概念ではない。例えば、大川（2002）はODAプロジェクトの評価をメタ評価し、妥当性、論理性、費用効果の視点の欠如などを指摘している。また国際協力機構外部有識者事業評価委員会/2次評価作業部会（2006）は、3回目の2次評価結果として公表されており、多くの事業評価報告書を客観的かつ定量的に評価し、共通する傾向や特徴的な傾向を導き出すことによって国際協力機構の評価の質の改善に寄与している。しかし、これらの研究は主として「評価が適切に行われたか」という評価にとどまる。つまり、事業評価として見た場合には、活動、あるいはプロセスの評価であって、アウトプット、アウトカムの評価ではない。評価事業を他の事業と同様に評価するためには、評価の目的である「改善を通じた質の向上」や「説明責任を果たす」という、アウトカムを評価しなければならないのは当然である。

本研究は、機構が過去に行った試行的大学評価を1つの事業としてとらえ、そのアウトプットやアウトカムの評価を行うことにより、優れた点と問題点とを明らかにするとともに、試行的評価で得られた経験と教訓とを新たな評価システムに活かすことを目的とする。

2. 大学評価・学位授与機構による試行的大学評価の概要

(1) 試行的大学評価の背景

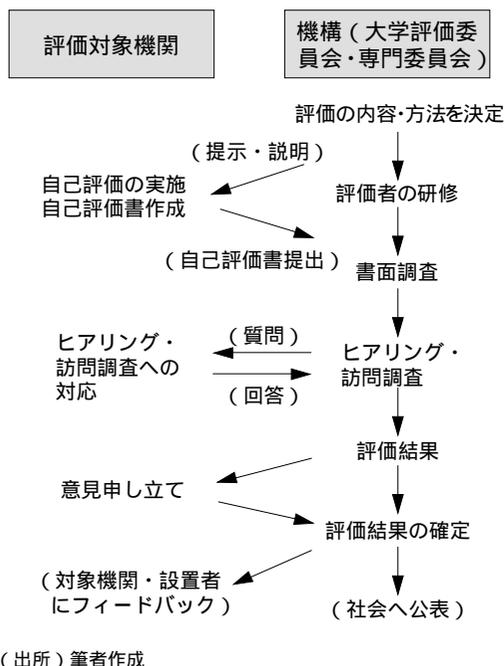
大学の評価は、1991年の大学設置基準の大綱化に伴い、自己点検・評価が努力義務化されたことで大きく展開した。つまり、大学に対する法的規制を弾力化して大学の自主的かつ自律的な改革を

図る一方、評価を通じて自らの努力で質の向上を行うよう促したのである。多くの大学が自己点検・評価を実施したものの、実際には形式的評価に陥り、改善に十分結び付かない例も多くあった。そのため、大学審議会（1998）は、自己点検・評価の充実を図るとともに、第三者評価システムの導入などを通じて多面的な評価を行い、大学の個性を伸ばし、教育研究の内容・方法の改善につなげるシステムを確立する必要を示した。これに基づき、1999年には自己点検・評価が努力義務から義務とされ、さらに外部評価が努力義務とされた。そして2000年には、大学の第三者評価を行う機関として大学評価・学位授与機構が改組によって設立された。

機構は、本格的な第三者評価の実施を2004年度に設定し、当面は大学評価に関する経験を蓄積するため、学問分野・機関数を絞って段階的に大学評価を実施することとした。この試行的評価は、2000年度から2003年度まで、設置者から要請のあった国立大学及び大学共同利用機関（以下「大学」と記す。）を対象とし、全学テーマ別評価、分野別教育評価、分野別研究評価の3つのテーマに分けて、3回のサイクルで実施された。全学テーマ別評価では評価対象は大学全体であり、主としてすべての対象機関の協力のもと、「教育サービス面における社会貢献」、「教養教育」、「研究活動面における社会との連携及び協力」、「国際的な連携及び交流活動」の4つの評価が実施された。分野別教育評価、同研究評価では学部または該当組織を対象に、理学、医学、法学、教育学、工学、人文学、経済学、農学、総合科学の9分野を対象とする評価が実施された。4年間の実施期間で評価を実施した機関は122機関、延べ550組織に上った。

2004年度から予定されていた第三者評価の本格実施は、学校教育法の改正及び国立大学法人法の制定等により実際には行われなかった。しかし、機構は、大学、短期大学、法科大学院、高等専門学校を対象とした認証評価機関の一つとして2005年度から認証評価を実施しており、また国立大学法人評価に際しても国立大学法人評価委員会の要請を受けて教育・研究活動に関する評価を実施する予定である。試行的大学評価で得られた教訓は、これらの評価の基礎として活かされる。

図1 大学評価のプロセス



(2) 試行的大学評価の方法

機構による試行的大学評価は、「大学の教育研究水準の向上に資するため、大学の教育研究活動等の状況について評価を行い、その結果について、当該大学及びその設置者に提供し、並びに公表すること」を目的とし、実施された。

図1は機構の行った試行的評価のプロセスを示したものである。評価は評価対象機関による自己評価と、訪問調査あるいはヒアリングが基礎となる。最初に、機構において大学評価委員会、各専門委員会が評価の内容・方法を決定し、評価対象機関に提示、説明を行う。評価対象機関はそれに基づいて自己評価を行う。専門委員会は提出された自己評価書に関して書面調査を行い、不明な点等についてヒアリングまたは訪問調査で確認し、そして評価結果が作成される。なお、評価に際しては、機関別に編成された評価チームや、学問分野別に編成された部会が実際の評価にあたる。評価結果は公表の前に各評価対象機関に通知され、意見申し立てを経て評価結果が確定される。最終

的に、評価報告書は対象機関・設置者にフィードバックされることにより、対象機関における質的向上を促進することが期待され、また社会に対する公表によって、アカウンタビリティの促進が期待されている。

3. 分析方法

本研究で分析対象としたデータは、試行的評価の中であらゆる機会で作成してきた(1)意見調査の結果に加え、今回の検証作業で改めて(2)質問紙調査、(3)インタビュー調査を実施し、収集したものである。

(1)意見調査：機構が評価実施の際に、各実施段階において評価対象機関、評価者、関係団体に行った意見照会の結果である。調査方法は、主として自由記述形式の質問紙を用いての郵送調査であり、文章単位で5,000以上の意見が収集された。それらの意見を取りまとめた結果、およそ500の意見に集約された。なお、詳細な結果は紙面の都合で省略しているが、研究評価については林ら(2003)、全学テーマ別評価については齋藤(2004, 2005)、ならびに試行的評価に関する検証委員会(2004)を参照されたい。

(2)質問紙調査：質問紙は(1)の集約された意見を参考に作成され、調査は平成16年7月にすべての評価対象機関に対して郵送法で行われた。この調査では、(1)の調査でカバーされていない、大学内の評価の過程や評価の成果等についての情報収集を行った。質問紙への回答は、各機関内での自己評価を実際に担当した方にお願ひし、回収数は539(回収率98%)であった。

(3)インタビュー調査：地域や規模等を考慮し、9大学及び1大学共同利用機関に協力をお願いし、平成16年9月から10月にかけてインタビュー調査を実施した。調査協力者は、評価担当者及び大学の運営に携わる方々であり、(1)(2)で十分に把握できなかった大学内の事柄について調査を行った。

(2)(3)の調査及び検証作業は、機構の試行的評価の当初の実施計画に基づき設計されたロジックモデルに基づいて行われた。ロジックモデルとは、プログラムの要素(資源)活動、アウトプット、アウトカムについて、またそれらの因果関係について明確なモデルとして示し、プログラムの問題解決に役立てる手法である(Bickman 1987, Wholey 1983, 1987)。実用に際しては、左から右へ、因果関係を同定した図を用いることが多い(例えばW.K. Kellogg Foundation 1998, 2004)。ロジックモデルは、問題解決にどの点を重視するかによって、セオリー・モデル、アウトカム・モデル、アクティビティ・モデル(あるいはそれらの複合モデル)と分類されたり(W.K. Kellogg Foundation 2004)。モデルの使用方法に関しても、プログラムのライフスパンの中で、フィージビリティ評価のため、改善のため、モニタリングシステムの開発のため、そして教訓の蓄積のためと、様々な目的で使用されたりする(Savaya and Waysman 2005)。また、ロジックモデルという名称や形式以外にも、プログラム・セオリーを使用するという点を中心に、多くのバリエーションが存在している(McLaughlin and Jordan 1999)。どの方法においても、プログラムを因果関係で捉え、問題点を抽出し、プログラムをより良いものにするという点は変わらない。そして、ロジックモデルの形式は変更不可能な決まり事ではなく、評価対象となる事業を明確に表す上での適切なモデルの選択は評価者に委ねられる。

本研究では、科学技術政策評価で定評のある英国マンチェスター大学のPREST(Policy Research in Engineering, Science and Technology: 工学科学技術政策研究所)の提唱するロジックダイアグラムを基礎として、プログラムのモデル構築を行った。このロジックダイアグラムでは、上下の階層構造で上位レベルの目的、下位レベルの目的、活動、期待されるアウトプット、アウトカム、インパクトの順に要素を並べ、それらの因果関係をリンクさせる。目的と期待されるアウトカムを別の階層とすることにより、目的には明示されにくい効果、負の効果、スピルオーバー的な効果についても明確に示されることを特徴とする(林 2002)。本研究でロジックダイアグラムを用いる目的は、試行

的大学の評価を一つの事業として捉え、今後の大学評価のための論点整理とともに教訓を導くことにある。その点、PRESTのモデルは主たるプログラムの目的のみならず、さまざまな成果を考慮することができる点でふさわしいと考えられる。なお、ロジックダイアグラムの構築に関しては、試行的評価が当初の評価設計通りに遂行されたか、またその遂行に際して当初の想定と異なる点がないかどうか等を検討することが重要であるとの考えのもと、当初の評価の設計(基本的枠組み)をそのままモデルに組み込んで分析を行った。

以下に示すような45の要因を設定し、因果関係に基づき構築されたロジックダイアグラムは後の図6に示すとおりである。

- ▶ 目的(2): 大学の質の向上とアカウンタビリティ
- ▶ 目標(5): 目的を達成するための具体的な課題(必須とする要素)
- ▶ 基本的枠組み(9): 評価手法として容れなければならない要因
- ▶ 投入(評価の体制)(機構3、対象機関1): 評価、改善、公表の各体制
- ▶ 活動(評価の実施)(機構4、対象機関1): 評価に伴う各種の活動(評価方法を含む)
- ▶ 結果(5): 評価結果、評価経験の蓄積と改善、作業負担、マスメディアによる報道
- ▶ 成果(対象機関8、社会5): 評価経験の蓄積、活動の改善、社会からの理解等の各種成果
- ▶ 目的の達成(2): 大学の質の向上とアカウンタビリティ

4. 分析結果

分析は、試行的評価のインプット、プロセス、アウトプット、アウトカムのすべてについて、網羅的に実施した。その結果、「目的」、「目標」の構成、一連の「評価の基本的な枠組み」に関してはおおむね良好な結果であり、一部の要素を除いて設計通り実施されたという結論を得た。評価を行う機構側の「活動」の問題点として、評価の目的や方法の理解に関する問題等が挙げられたが、試行的評価の進行とともに改善された。一方、評

価の対象機関側の「投入」と「活動」(図6:中央右側)一連の「結果」及び「成果」(図6:下段)は、その一部または全体的に、問題が生じていたことが明らかとなった。

そのため、ここでは対象機関側の「投入」と「活動」、つまり自己評価のプロセスと、試行的評価全体としての「結果」と「成果」の分析結果について論ずる。

(1) 自己評価のプロセス(対象機関側の「投入」と「活動」)

質問紙調査の結果、多くの機関では自己評価を担当する組織として委員会形式の組織が挙げられた。これら組織は約9割が常設の組織であり、既存の自己点検・評価の体制で試行的評価に対応していたことが明らかとなった。そして、実際の作業のために、評価担当組織の下にワーキンググループや部会などを設けている機関が多くを占めた。インタビュー調査では、当初は委員会組織で対応していたものの、実際には対象テーマに関する専門性が必要となることからワーキンググループ等を設置した機関もあり、過年度の評価経験を踏まえて適切な自己評価体制を模索した過程がうかがえる。また、中には評価に特化した常設の専門組織を設置している機関も見られた。インタビューでは、評価専門部署を設けることで、より全学的な組織的連携の下で評価を実行、統括できるため、問題点の把握等も認識しやすくなったとのことであり、この試行的評価の経験が昨今の国立大学内での評価関係組織の充実に繋がっている。

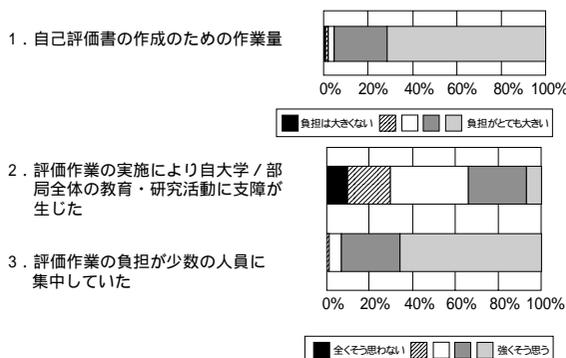
自己評価等の作業量については、意見調査では評価担当者に過大な負担が強いられているという意見がきわめて多くあった。図2は自己評価担当者の作業負担に関する質問紙調査の結果であるが、「1.自己評価書作成のための作業量」について、「負担が大きい(5段階評価のうち4と5)」とする回答が9割を超えており、評価負担の軽減は必要不可欠と言える。しかし、「3.評価作業の負担が少数の人員に集中していた」という質問に対する肯定的意見も同じく9割を超えていることから、負担が大きい理由として、評価作業が日常の

業務に位置づけられていないこと、そして評価作業が一部の部署・人員に集中しているという評価対象組織内の要因が挙げられる。

(2) 評価結果の適切さと公表方法(「結果」)

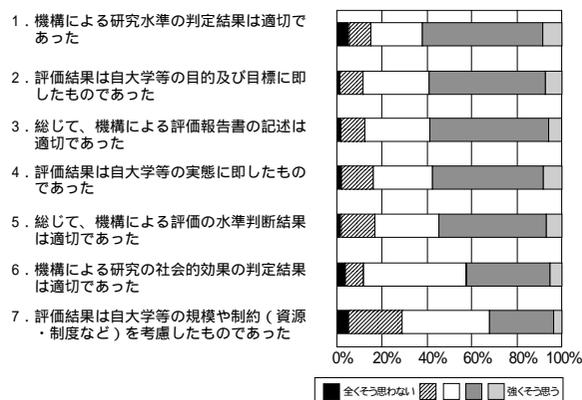
試行的評価の評価結果は、大学の目的に即して、評価項目ごとに判断結果を定型句で表現するとともに、優れた点と改善を要する点を評価報告書の中で指摘した。評価結果は試行的評価事業の評価の中で最も配慮すべき事柄の一つであるが、この評価結果(評価報告書の内容)の適切性について、評価対象機関の回答を表したものが図3である。

図2 自己評価担当者の作業負担



(出所)筆者作成

図3 評価結果の適切性



(出所)筆者作成

この結果から、評価結果は評価対象機関からおおむね肯定的に受け入れられている。判定結果、報告書の記述とも大学からの評価は悪くはない。しかし、「6.研究の社会的効果」、「7.自大学の規模や制約（資源・制度など）の考慮」については肯定的回答の割合が相対的に低く、比較的困難な社会的効果の慎重な扱い、そして各大学の特性をより考慮した評価が求められている。

これらの評価結果が「成果」を生むためには、それが関係者に伝わらなければならない。対象機関に対しては直接通知するために問題はないが、社会に対しては評価機関が結果を公開するだけでは周知に限界があるため、ロジックダイアグラム上、評価の「活動」と社会における「成果」とを結ぶ「マスメディアにおける報道」の役割が重要となる。

しかし、意見調査において、評価結果の公表方法に対する非常に多くの批判が寄せられた。特にマスメディア等による評価結果の取扱いについての批判である。評価結果は目的に即したものであるため、結果の比較は意味を持たないという認識を評価者側と評価対象機関で共有し、評価を実施し、記者発表の際にも十分に説明をしたにもかかわらず、評価結果がランキングの形で報道されたり、大学間の比較をされたり、あるいは良い結果を得た項目は報じられず悪い結果のみがクローズアップされるなど、大学評価の意図とその報道のされ方には大きな隔たりがあった。つまり、そのような評価結果の表し方は、現状の報道には適さないということであり、さらには社会の捉え方とも乖離しているということである。わかりやすい評価という意味では、その要求に答えられていなかった点を見過ごすことはできないし、逆に社会にランキングではない大学評価手法の説明をする必要もある。メディアが大学評価に果たす役割は、機構の評価が最終的に判断を社会に委ねる以上、機構と大学の利害関係者とを結ぶ意味で極めて重要であり、この関係は新たな評価制度でも変わらない。評価の趣旨・ねらいやその評価結果のとらえ方等、さらには大学の実際の状況の理解について、社会からよりの確な理解が得られるよう、社会に対する、機構、大学双方の一層の努力が求められる。

(3)大学の質の向上(対象機関で発現する「成果」)

大学の質の向上を実現するために、試行的評価の基本的枠組みでは「大学の主体的な改善を促す評価」、「大学の個性を伸ばす評価」が目標とされた。これらの目標の達成について質問紙調査で確認したところ、その結果は図4のようになった。「2.当該活動の改善を促進した」、「5.組織的に運営する重要性が自大学内に浸透した」等、6割以上が肯定的な回答であり、自己点検・評価との比較を求めた項目も肯定的回答である。当初の設計通り、評価が改善を通じての質の向上に繋がっていることが見て取れる。

これは、自己評価よりも第三者評価が優れているということの意味するものではない。インタビュー調査では、自己評価によって自大学の課題を把握することができたというような意見が多くあり、評価の一環として、自組織に都合の良い自己点検・評価ではなく、外部の者によって決められた項目で実施した自己評価作業自体が、大学の改善に向けて有効に機能したものと考えられる。また、大学評価が、その後の中期計画の策定、マネジメントの改善や、評価の重要性の認識等にも繋がっていることは、その効果が多様であることを意味している。

さらに、評価後の具体的な改善事例についての記述を求め、それらの各改善事例について機構の評価がどの程度参考となったかを質問したところ、全対象機関122機関中の109機関(89%)から、のべ1,024件の回答が得られた。その結果をまとめたものが表1である。

回答のあった改善事例のうち、全評価区分を通じて多かった事例としては、学内組織の連携強化/整備に関するものや、自大学内での評価体制の整備や改善システムの強化(体制整備、根拠資料等の収集の経常化、アンケート等による問題点の把握など)、周知公表の体制(活動)の強化・充実など、教育研究活動等を実施する上での基盤的な部分の更なる改善に関する事項が挙げられる。これらについて、評価の参考度が4(ある程度参考となった)以上であるものは825件(81%)となった。このことから、改善の促進に関して試行的評価の一定程度の貢献が認められる。

図4 大学の質の向上



(出所) 筆者作成

しかし、これら改善の促進は、評価結果を受けての改善のみならず、試行的評価への組織としての対応、また自己評価の過程が直接的に改善を促したとする意見がインタビュー調査では多くあった。例えば、既に自己点検・評価や日常の活動の中で認識されていた課題に関して、試行的評価を契機に積年の課題への取り組みが一挙に進んだというコメントもあり、機構の評価が改善の促進装置として働いたことが指摘できる。これは当初の想定とは異なる因果関係となっている。さらに、「評価対象の活動を組織的に運営することの重要

性が自大学内に浸透した」などのマネジメント上の意識改革の効果についても半数以上の肯定的回答が得られており、大学の改善への間接的な効果も少なからず存在したと言える。

一方、個性の伸長については肯定的な回答が4割程度にとどまった。このことは、大学の多様性を尊重するために採られた「目的に即した評価」という評価の枠組みが、個性を犠牲にしないためには機能しても、個性の促進のためには十分ではないことを示している。また、試行的評価を実行可能なものとするために、項目、要素、観点等の

表1 改善事例の件数と評価の参考度

評価区分	改善事例	1	2	3	4	5	記述なし 参考度の 総計	総計	(回答者数)
		全く参考と ならなかった	~	言えない どちらとも	~	非常に参考 となった			
全学テーマ 別評価	テーマの実施・改善のための体制	6	13	66	235	100	7	427	425
	実施方法や実施内容	2	11	35	158	71	9	286	
分野別 教育評価	教育の実施・改善のための体制	1	1	11	44	16	2	75	65
	教育内容・方法や学習支援の方法	1		10	47	19	1	78	
分野別 研究評価	研究実施・支援・改善のための体制	1	2	5	45	21	4	78	62
	研究活動のための諸施策、諸機能			8	36	11		55	
	研究内容		1	2	16	6		25	
総計		11	28	137	581	244	23	1,024	552

教育評価と研究評価にはともに「総合科学」を含む。

(出所) 筆者作成

設定をしたことなどもこの結果に影響を与えていると考えられる。

問題点としては、自己評価に携わった担当者間では、評価の目的や評価に基づく改善等の枠組みを十分に理解しているが、それが少数の人員に限られており、評価に携わった者が有する情報や経験が他の教職員には十分伝わっていないという状況にあることが各調査から明らかとなった。組織としてのさらなる質の向上のためには、情報や経験が組織の中で広く伝わるような体制の確立が不可欠であると考えられる。

(4) 大学の諸活動の社会への説明(アカウンタビリティ)に関する成果(社会で発現する「成果」)

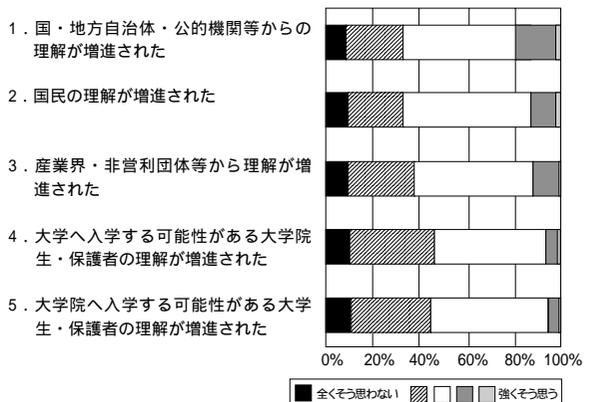
試行的評価の2つ目の目的である「大学の諸活動の社会への説明」の実現のためには、評価自体が「社会が大学の状況を把握できる評価」であることが求められた。機構の評価結果によって社会の理解が増進されたかについて質問した結果が図5である。

回答の平均値は2.5~2.8と他の質問項目に比して著しく低く、全体の3~4割が否定的な回答であった。機構では評価結果を対象機関及び設置者に

提供するとともに、マスメディアへの記者発表、印刷物の刊行、ウェブサイトの掲載等により広く社会に公表してきたが、評価委員からは、マスメディアによる報道を受けて、逆に大学の活動について誤解を生じる可能性があるというコメントすら寄せられている。

特に着目すべきは、入学希望者の活用の程度である。その他の利害関係者と比較しても低い値となっており、最も活用してもらいたい者に活用されていない状況は改善の余地がある。評価報告書

図5 関係者による大学の理解の促進



(出所) 筆者作成

を高校生やその保護者、産業界、国民一般が読むには、その内容が難しく、また分量が多すぎて、これらの人々が大学について知りたい情報が十分含まれていると言える状態にはなっていないという指摘もあった。評価担当者からも「社会に広く読まれるような分かりやすい内容と形式にすべき」との意見も多く、評価報告書の内容、あるいは評価報告書以外の公表の方法について、改善の余地が残されている。

(5) 総合的分析結果

分析結果を総合的に判断し、それをロジックダイアグラム上に表現したものが図6である。これまで分析してきた、ダイアグラムの中の「投入」から下の階層に関して、問題がなく実施されたと評価される要素は実線で囲んだままであるが、やや問題が生じたと評価される要素は点線で囲んで印を、そして問題があると評価される要素は線で囲まれずに×印を付した。

一番下の層で示されているように、目的の一つである「大学の教育研究活動等の改善」については、大学における改善にある程度の貢献が認められること、大学のマネジメントや教職員の意識改革への影響、及び大学の質の向上に関する間接的な効果も存在することなどから、まだ十分な成果の発現には時間を必要とするものの、おおむね達成できていると考えられる。その成果の発現を阻害する要因としては、ロジックダイアグラムに沿った分析からは、大学の中での評価経験の蓄積、評価の趣旨・制度等の浸透が挙げられる。前述のように、評価に携わる人間が一部にとどまり、またそこからの情報共有が進まないという状況は設計時には想定されていなかった。また、大学からは、各大学の個性の促進についてより一層考慮することが望まれており、評価側にも改善すべき点が多く残されている。

一方、もう一つの目的である「広く国民の理解と支持が得られるよう支援・促進」については、利害関係者が評価結果を活用したという事例が十分には確認できず、大学評価の仕組みや評価結果に関する社会全般の理解度や活用のされ方という点での課題が多く、達成されていない。この問題

は、ロジックダイアグラム上では、評価結果の公表方法や、マスメディアによる報道がネックとして考えられる。これについては、そもそも「社会は大学に関する情報を欲している」という前提が形成されていたことも1つの原因と考えられる。つまり、社会は大学のどんな情報でも良いというのではなく、わかりやすい情報、必要とする情報を求めていると言えるが、現在のニーズからそれをランキングという形に求めるのはあまりに安易である。第三者評価機関でなくては提供し得ない、しかも利害関係者が真に必要とする情報について、高い質を保ち、必要な形で提供することが求められている。

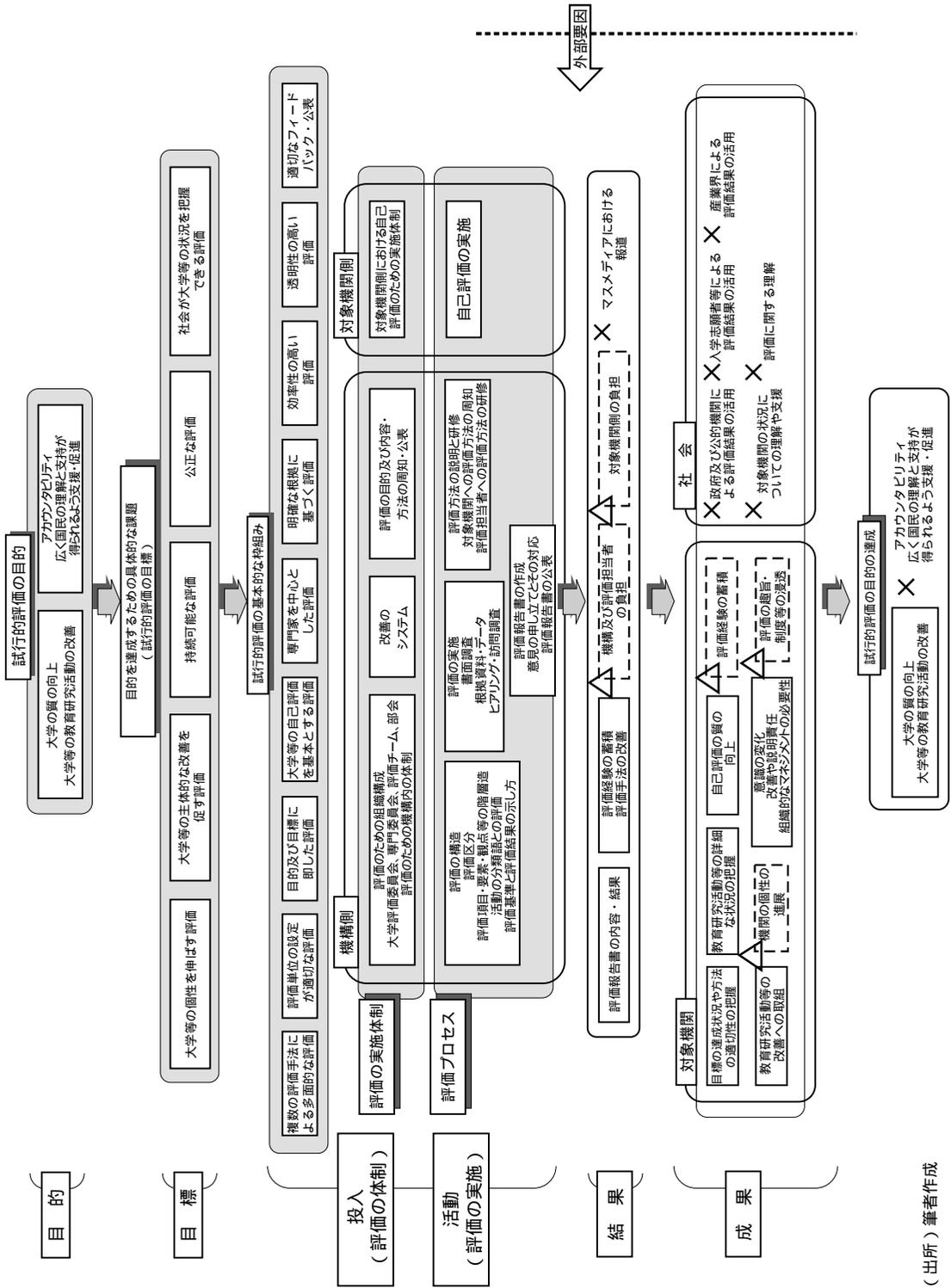
5. 作業負担の検討

本研究では当該評価の検証を行うためのロジックダイアグラムを設定する際に、あえて投入としての資源を考慮せず、作業負担という形でアウトプットに配置して分析を行った。これは、機構による試行的大学評価が今後の大学評価の基盤を作るものであるために、投入資源は所与のものとしてほとんど無視できたためである。しかし、実際には資源は無限ではなく、評価側のみならず、評価を受ける側の資源についても検討を行わなければならない。作業負担が大きいという結果は、つまりは当初の投入資源の見積もり、あるいは評価設計自体に無理があったということである。

評価に携わる者は、評価を受ける側も行う側も、社会的に見ても能力が高く、評価に費やした時間を他の業務に振り向ければどの程度の高い生産性があるのかについても本来は検討すべきであろう。大学評価は、その生産性以上に効果がある事業であることが求められるし、また評価の規模が大きくなればなるほど、評価に費やされる資源も増大する。そこで、ここでは今後の大学評価設計の参考として、評価担当者が評価に要した時間を算出し、検討を行う。

評価担当者は、2年間の評価期間中に、書面調査等の評価作業に携わるだけでなく、各種の会議に出席し、またヒアリングまたは訪問調査に参加する。これらの評価活動に要した時間には、2001

図6 試行的大学評価のロジックダイアグラム



(出所) 筆者作成

年度から2002年度にかけての2サイクル目に行った評価の中で、全学テーマの「研究活動面における社会との連携及び協力」、分野別教育評価の「工学系」、分野別研究評価の「教育系」および「法学系」のデータが利用可能であることから、それらを使用した。評価担当者による質問紙への回答であることから、あくまでも推測値である。

表2の結果を見ると、評価担当者がどの評価種別においても平均しておよそ150時間以上を評価に費やしていることがわかる。日数に換算するとおよそ1ヶ月であり、評価の主査、副主査を担当している場合にはもっと多くの時間を費やしていることを考えると、負担は極めて大きいと言わざるをえない。その主な原因は、研究成果の査読に要した時間である。試行的評価に関しては各研究者の研究成果を一定数出してもらい、それらをすべてピア（当該分野を専門とする他の研究者）が評価したことから、このような結果となっている。

対象機関あたりに要した時間に換算すると、全学テーマ別評価では対象機関の数が多いために91

時間と一番少なく、そして分野別研究評価（教育系）がおよそ3,122時間という結果になった。評価担当者が評価に費やした時間だけでこれだけの大きな負担であることから、評価対象機関の負担も加えればより多くの資源が用いられたこととなる。

これらの値からは、例えば分野別研究評価を受けた法学部では、少なくとも1,865時間の労力に相当する効果をこの評価から得られなければ、社会的に見て効果が生じていないこととなる。しかし、たとえば評価の効果が長期間持続すれば、その値は割り引かれるはずである。大学が定期的に受けることが義務づけられている認証評価は、その期間が7年間であり、そして国立大学法人が受ける国立大学法人評価の中期目標期間は6年間である。ここで単純に、対象機関あたり平均総作業時間を6年間で割れば、表2の全学テーマ（研究連携）分野別教育（工学系）分野別研究（教育系）同（法学系）の順に、それぞれ15.1時間、168時間、520時間、311時間となり、それだけの費やされた時間を1年間で取りかえさなければならない

表2 評価担当者の総機会費用

費目分類		評価種別	全学テーマ (研究連携)	分野別教育 (工学系)	分野別研究 (教育系)	分野別研究 (法学系)
(固定) 会議・打合せ			1,462	2,177	1,714	1,605
機関別 チーム	(変動) 会議・打合せ		1,915	792	525	503
	(変動) 評価作業		4,246	1,488	1,728	1,632
研究分野 別部会	(変動) 会議・打合せ				2,248	1,720
	(変動) 評価作業（査読）				12,008	5,348
(変動) ヒアリング・訪問調査			2,644	1,608	510	381
合計			10,266	6,065	18,732	11,189
評価者あたり平均作業時間			168時間	202時間	226時間	145時間
対象機関あたり平均総作業時間			91時間	1,011時間	3,122時間	1,865時間

((断りがない限り) 単位: 人・時間)

(注) 費目分類は次のように対応する。なお、(固定)(変動)とは、評価規模による増減の有無を表す。

(固定) 会議・打合せ: 専門委員会、主査打合せ等

(変動) 会議・打合せ: チーム・部会会議、説明会・研修会等

(変動) 評価作業: 書面調査、査読作業、判定等

機関別チームとは評価対象機関別に構成されたグループを、研究分野別部会とは分野（教育学、法学等）の中のにさらに細かい研究分野別に構成されたグループを示す。

会議・打合せに関しては、本務からの移動時間を含めていることから、実質的な作業時間は表中の値の1/3程度である。

評価作業は質問紙調査を元にした平均値で算出しており、中央値で算出すると3/4程度である。

(出所) 筆者作成

ことになる。

ある程度予測可能な範囲内とも考えられるが、研究評価は大きな負担を評価者に強いてしまったこととなる。教育評価と比較しても、研究評価の特に査読に相当する作業は軽減される必要がある。

6. まとめと提言

機構の試行的大学評価について詳細に評価を行った結果、目的の一つである大学の教育研究活動等の改善については、おおむね達成できていると評価された。一方、もう一つの目的であるアカウンタビリティの促進については、課題が未だ多く、改善の余地があると評価された。

大学の質の向上については、大学内で一部の評価担当者に作業が集中する弊害と、評価担当者の経験と情報の蓄積を組織全体に浸透させる段階が機能していないことが示された。機構としてはこのような評価に関する知識、経験の波及について、大学により積極的に働きかけていく必要がある。

また、社会からの理解については、機構や大学と、社会とを結ぶ重要な役割を担うマスメディアに重要な役割を担ってもらわなければならないことが、ロジックダイアグラム上からも明らかであり、マスメディアに対して評価の趣旨を十分に理解してもらうよう努力することが必要である。また、一方で、社会に対してそれとは別のロジック（因果関係）を用意すること、つまり別のチャンネルを用意することも重要であろう。インタビュー調査では、大学自らが情報を積極的に発信している大学は、新聞等の報道をさほど問題視していなかった。このことから、大学自らの努力で社会の理解と指示を得られるようにすることもまた必要であると考えられる。

機構の新しい大学評価制度においても、社会への因果関係をいかに成立させるかという点が懸案事項であることに変わりはなく、広報活動の充実等によって大学や社会に積極的な働きかけを行い、また評価システムを改善していくことで対応する必要がある。

機構の評価は大学の質の改善とアカウンタビリティを目的としつつも、機構の役割はそれらの支

援でしかなく、機構は評価結果を公表することはできても、そこから先の大学の質の改善とアカウンタビリティの促進は大学の手にな任せられ、さらに最終的な評価結果の活用は利害関係者に委ねられる。すなわち、ロジックダイアグラムの因果関係の先の方、プログラムの実施者の手が届くか届かないかという先に、成果の成就を決めるほとんどの因果関係が存在するという状況にある。

認証評価の制度の中で、多くの認証評価機関は大学の連合体という背景を有する一方で、機構は独立行政法人という設置形態である。このことが、大学の自律性との関係で、余計に協力関係を構築しづらい、手を出そうにも出せない枷となっている。

機構は、大学評価に関してはあくまでも第三者として、評価対象機関と利害関係者にどのような情報を提供していくべきか、またどのような支援を行っていくべきか、誰に対しても直接的関与はしないにせよ、評価を通じて間接的に質的向上を適切に支援していくという、難しい立ち振る舞いが要求されている。作業負担や費用効果等の評価の実施可能性とともに、大学評価の成果の発現までを見越した評価システムの設計と、因果関係が確実に成立し、成果が発現するような、すべての評価の利害関係者等への間接的支援が必要である。

謝辞

本研究は意見照会並びに各種調査に応じていただいた評価対象機関、評価者、関係団体の方々の回答を基に分析を行ったものである。ここに厚く御礼申し上げます。特に意見集計については、大学評価・学位授与機構の川口昭彦教授、芳鐘冬樹助手、野澤孝之助手にご協力いただき、また有意義なコメントをいただいた。ここに厚く御礼申し上げます。

参考文献

- Bickman, L. (1987). The Functions of Program Theory. ” In L. Bickman (ed.), *Using Program Theory in Evaluation*. New Directions for Program Evaluation, 33, San Francisco, Jossey-Bass.

- 大学審議会（1998）「21世紀の大学像と今後の改革方策について - 競争的環境の中で個性が輝く大学 - 」
- 林隆之・齋藤貴浩・米澤彰純・川口昭彦（2003）「大学評価システムのメタ評価：大学評価・学位授与機構による研究評価の運営レビュー」、『研究・技術計画学会第18回年次学術大会講演論文集』、510-513
- 林隆之（2002）「欧州における研究評価の教育コースの現状」、『大学評価』、第2号、大学評価・学位授与機構、35-53
- 国際協力機構外部有識者事業評価委員会 / 2次評価作業部会（2006）「2005年度2次評価結果」、第1章、「外部有識者事業評価委員会による2次評価」、第4部、『事業評価年次報告書2005』、国際協力機構、114-140
- 国立大学協会（2002）『平成12年度着手の大学評価に対する意見について』（http://www.janu.jp/active/txt5/h14_7_2_a.html）（2006.11.1取得）
- 国立大学協会（2003a）『平成13年度着手の大学評価に関する申し入れ』（http://www.janu.jp/active/txt5/h15_6_6.html）（2006.11.1取得）
- 国立大学協会（2003b）『「大学評価・学位授与機構の評価事業の今後の在り方について（中間まとめ）」について（意見）』（http://www.janu.jp/active/txt5/h15_9_25.html）（2006.11.1取得）
- McLaughlin, J. A., and G. B. Jordan（1999）. Logic models: A tool for telling your program's performance story. *Evaluation and Program Planning*, 22, 65-72.
- 大川夏子（2002）「プロジェクト評価のメタ評価 - 人口・家族計画プロジェクトを例として - 」、『日本評価研究』、2（2）、91-100
- 齋藤貴浩（2004）「大学評価・学位授与機構による大学評価の試行に対する意見集計結果 - 全学テーマを対象として - 」、『日本評価学会春季第1回全国大会発表要旨集録』、167-174
- 齋藤貴浩（2005）「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」、『日本評価学会第6回全国大会発表要旨集録』、257-260
- Savaya, R and Waysman, M.（2005）“The Logic Model: A Tool for Incorporating Theory in Development and Evaluation of Programs.” *Administration in Social Work*, 29, 2, 85-103.
- 試行的評価に関する検証委員会（2004）『大学評価・学位授与機構が平成12年度から平成15年度までに実施した試行的評価に関する検証について - 試行的評価に関する検証結果報告書 - 』、大学評価・学位授与機構
- Wholey, J. S.（1983）. *Evaluation and Effective Public Management*. Boston, Little, Brown.
- Wholey, J. S.（1987）. Evaluability Assessment: Developing Program Theory. In L. Bickman (ed.), *Using Program Theory in Evaluation*. New Directions for Program Evaluation, 33, San Francisco, Jossey-Bass.
- W.K. Kellogg Foundation（1998）. *The W.K. Foundation Evaluation Handbook*. Battle Creek, W.K. Kellogg Foundation. (<http://www.wkcf.org/Pubs/Tools/Evaluation/Pub770.pdf>)（2006.11.2取得）
- W.K. Kellogg Foundation（2004）. *The W.K. Foundation Logic Model Development Guide*. Battle Creek, W.K. Kellogg Foundation. (<http://www.wkcf.org/Pubs/Tools/Evaluation/Pub3669.pdf>)（2006.11.2取得）
- （2007.2.19受理）

**For Further Improvement of Universities:
Evaluation of the Evaluation Program
Implemented by National Institution for Academic Degrees and University Evaluation
in Japan**

Takahiro Saito Takayuki Hayashi

Faculty of University Evaluation and Research
National Institution for Academic Degrees and University Evaluation
tksaito@niad.ac.jp

Abstract

From academic years 2000 to 2003, the National Institution for Academic Degrees and University Evaluation (NIAD-UE) implemented evaluations of selected universities on a trial basis. Evaluation can be regarded as an intervention and subsequently the target of a program evaluation. After the NIAD-UE trial evaluation was completed, a meta-evaluation was conducted to clarify its outputs and outcomes, including successful aspects and problems, through questionnaire surveys and interviews with the evaluators and the evaluated. As a result, it was made clear that the process and the result of the evaluation had led to the improvement of universities' educational and research activities. However, many of the persons engaged in the evaluation pointed out their considerable burden. Apathy of both other university staff and the public about the evaluation also remained a problem for making an effective evaluation system. Moreover, the concept of cost effectiveness should be introduced to the new settings of university evaluations, because the evaluation cost was worse especially for research assessment.

Keywords

university evaluation, meta-evaluation, logic diagram (logic model),
cost effectiveness, self-evaluation

【研究論文：依頼原稿】

開発途上国における教育評価に関する理論的比較研究 - 国際学力調査、学校調査、世帯調査の視点 -

西村 幹子

神戸大学大学院国際協力研究科
mnishimu@kobe-u.ac.jp

要 約

本稿は、既存の教育評価手法が、開発途上国において有効な教育政策や教育計画の策定に繋がるためにいかなる課題を有しているかを、手法の理論的比較分析という視点から考察し、認識枠組みとしての総合的教育評価モデルを提案するものである。具体的には、教育評価についての既存の国際学力調査、学校調査、世帯調査を比較・対比させ、課題の抽出を行った。考察結果として、教育評価に使用されるデータの入手方法において、従来の国際学力調査がその対象や実施方法において低所得国への示唆を得ることが困難であること、学校調査や世帯調査も、それぞれ教育を供給する側と需要する側の片方に視点を当てることにより、教育開発に関する包括的な方策を提示するに至っていないことが確認できる。これらの課題は先進国にも共通するものであるが、開発途上国においては各種調査の結果が大きく乖離することがあり、総合の必要性は特に大きい。また、定量的なデータだけでは得られない推論（inference）を定性的な調査から得ることも、非就学や退学などの問題に対するきめ細かい対応には不可欠となる。

キーワード

教育評価、開発途上国、定量データ、学校調査、世帯調査

1. はじめに

近年、開発途上国における教育計画は画一的になっている、との批判がある（Samoff 1999; Foster 2000; Brown, et al. 2001）。これは、各国の教育計画が援助国と被援助国の一部の政策エリートによって作られているという点と、実際には被援助国政府には実質的な影響力がなく、援助国側の計画立案体制の押し付けになっているとの指摘によるものである。また、これらの指摘は、教育計画と並んで教育評価が画一的となっていること

をも示唆している。確かに、ミレニアム開発目標（MDGs）などの教育関連の国際目標には、就学率、修了率、識字率、男女間格差の是正など、一般的かつ定量的な指標が並んでいる。また、万人のための教育（EFA）・ダカール行動枠組みには、教育の質の向上についても言及があるが、EFAモニタリングレポートに掲載されているEFA開発指標（EDI）の中で質を示す指標としては、多国間の比較可能性の制約から初等教育5学年までの達成率が用いられている（UNESCO 2006）。

ニューパブリック・マネージメントなどの成果主義を重視する流れの中で、開発途上国で作成されている貧困削減戦略ペーパー（PRSP）も一般的な教育の現状には触れているが、それらがなぜ起こっているのか、という個別の教育政策を導くような分析は殆ど行われていないことが報告されている（World Bank 2005; UNESCO 2006）。指標の取られ方についての理論的考察が不十分である点についてはいくつかの先行文献があるが、いかなる調査方法が必要であるかについての具体的な提案はなされていない（e.g. Filmer, et al. 2006; UNESCO 2006）。

単に指標を収集し、その推移を追うだけでは教育計画や政策に結びつくような手がかりを見出すことはできない。例えば、どのような子どもたちが学校に行っていないのか、その理由は何か、を明確にしない限り、就学状況のみのデータを収集しても有効な政策にはつながらない。初等教育の学齢児童の約3～4割が就学しておらず、就学している生徒の約6～7割しか最終学年に到達しないような南アジアやサブサハラアフリカ地域などの状況を考慮すれば、学校データがいかに限定された層を対象としているかが分かる¹。MDGsが既にサブサハラアフリカ等の低所得国において達成不可能とされている状況下において（United Nations 2004）就学率や修了率などを上げるといふ国際目標を達成するためには、教育評価は、学校関連の指標を収集するだけでなく、教育と社会との関連に関する包括的なデータに基づいて行われなければならないだろう。

このような認識に基づき、本稿は、教育評価に使用される定量データの意味と、既存の学校調査と世帯調査の相違点に関する分析を通して、定量データの観点から、主に開発途上国における既存の教育評価指標の収集方法の問題点を明らかにすることを目的とする。それにより、教育評価の改善に資する検討を行うと共に、近年画一的になっているとされる開発途上国における教育計画と評価についての関係改善の糸口を見出すことを意図するものである。次節では、開発途上国における教育評価指標の意味について整理し、既存の教育評価に使用される定量データについて問題提起を行う。第3節では、学校調査と世帯調査を比較す

ることにより、教育評価に用いられる既存のデータ収集方法の有用性と限界を指摘する。第4節では、それらをふまえ、教育計画や政策に有効に還元する教育評価のための定性データを含むデータ収集のあり方について考察し、暫定的な結論を述べる。

2. 開発途上国における教育評価指標の意味

教育評価に関する研究者や教育活動に携わる実務者の多くは、教育の結果（outcome）として定量的なパフォーマンス指標を使用する。例えば、教育における費用効果分析は、ほぼ全ての場合において、教育の費用に対する教育の結果を測るために生徒のテストの成績を用いる（Levin 1995; Levin and McEwan 2001）。国際教育プロジェクトを計画するに際しても、経済協力開発機構（Organisation for Economic Cooperation and Development、以下、OECD）や国際教育達成度評価委員会（The International Association for the Evaluation of Educational Achievement 以下、IEA）による定量的なパフォーマンス指標が使用されてきた。テストの成績を費用効果分析や教育評価、計画に使用する理由は単純である。比較的入手するのが容易であり、図示しやすく、社会で求められる能力（competency）を代表するものと信じられているからである。つまり、教育の結果を測るために入手可能なツールとして成績を利用しているのである。

特に、国際比較を行うことの多い比較国際教育学の分野では、使用される成績の点数は読解（主にヨーロッパ言語）数学、理科などの各国の歴史的、文化的な影響が少なく、「比較可能」と考えられている科目に限定している場合が多い。例えば、IEAにより行われている国際数学・理科教育動向調査（Trends in International Mathematics and Science Study: TIMSS）は数学と理科に、読解力調査（Progress in International Reading Literacy Study: PIRLS）は読解力に限定して調査を行っている。また、東南部アフリカ諸国を対象に行われているSACMEQ（Southern and Eastern African Consortium for Monitoring Educational Quality）は

語圏アフリカで行われている教育システム分析プログラム (Programme de' analyse des systemes educatifs de la CONFEMEN: PASEC) およびラテンアメリカ諸国で実施されているラテンアメリカ教育質研究 (Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación: LLECE) は、読解 (英語、仏語、西語、ポルトガル語) と数学を対象としている。OECD加盟国による生徒の学習達成度調査 (Performance for International Student Assessment: PISA) は、問題解決領域を含むが、読解力、数学、科学での生徒の知識と技能を比較することに着眼点を置いている。

これらの国際的な定量データは、通常、教員の資格や経験、学校施設などの学校関連の指標や生徒の社会経済的な家庭環境に関する指標と共に収集される。このため、どのような生徒がどの程度の成績を取ったか、或いは家庭と学校のどのような要因が成績に有意に影響しているかを統計的な分析により知ることができる。このような方法で、開発途上国においては家庭の要因とともに学校の要因も数学や科学の成績に影響している、という結果を導き出したハイネマンらの業績 (Heynemann and Loxley 1983) に類似した研究を行うことができるようになった。

しかし、これらの入手可能なデータは、開発途上国における教育政策への反映という意味において、限られた科目に関するものであるという点以外に、2つの主要な限界を有している。一つは、調査の対象、もう一つはデータの入手方法である。

まず、調査の対象に関しては、参加国の制約がある。これらの既存の定量データは、PASECを除き低所得国の参加が非常に限定されている。国際的な研究で使用されるPISAやTIMSSにも、サブサハラアフリカ諸国などは殆ど参加しておらず、低所得国の参加は限定されている (表1参照)。実際、PISA およびLLECEには低所得国は1カ国も参加しておらず、TIMSSは46カ国中2カ国 (約4%)、PIRLSは47カ国中1カ国 (約2%)、SACMEQでも14カ国中8カ国 (約57%) である。2006年版の世界銀行の『世界開発報告書』に掲載されている209カ国中、59カ国 (約28%) は低所得国であることから、既存の学力調査がいかに低所得国の学力について知る手段として限定的であるかが分か

る。低所得国の参加が限定されている状態で、これらのデータを用い、国家横断的な分析をしたとしても、その結果を以って開発途上国全般への示唆を得ることは難しいという状況である。

また、調査の対象としては、初等教育レベルの高学年が採用される場合が殆どである点にも注目する必要がある。無論、学校教育による教育達成度を測るためには、ある程度の学習の蓄積が必要であり、数年の初等レベルの教育を以っておよその学力を測るという配慮は重要である。しかし、開発途上国においては、多くの子どもたちが学校に行っておらず、また入学したとしても退学することも珍しくないという状況がある。つまり、学校に行っていない、或いは学校に行ったとしても継続的な学習が困難であり、退学や留年を繰り返す子どもたちの学習環境や学習達成度については、学校を単位とし高学年を対象とした学力調査は限界を有しているといえる。

表2は、現在の開発途上国の教育指標を示したものである。全世界には初等教育の学齢に達しているにも拘わらず学校に行っていない児童が約1億人おり、そのうち開発途上国の児童が95%以上を占める。中でも、サブサハラアフリカおよび南西アジアが全体の7割を占めている。学齢児童のうち就学している児童は世界全体の約8割強であるが、南西アジアでは6割である。留年、退学などの教育の内部効率性も、全世界と開発途上国平均には大きな差があり、特にサブサハラアフリカや南西アジアの低所得国では3~4割が初等教育を修了せずに退学している。

こうした国々においては、例えば小学校6年生を標本として成績を測った場合、その結果が全体の小学生、あるいは学齢児童を代表しているとは考えにくい。仮に、これらの限定された学年の標本の成績に基づいて、社会経済的な家庭環境が生徒の成績に影響を及ぼす度合いが比較的小さい、という分析結果が得られたとしても、それは、あくまでもある学年まで到達した生徒の間での差に基づいて統計的に得られた示唆である。例えば、サブサハラアフリカであれば、全学齢児童のうち8割しか就学しておらず、そのうち修了する率は6割であるから、最終学年で学力を測った場合、全体として学齢児童の約半数の学力を以って教育の

表1 国際的な学力調査の参加国および対象学年

調査名(実施年)	参加国/地域	科目・対象学年/年齢
PISA(2000年および2003年)	アイスランド、アイルランド、アメリカ、イギリス、イタリア、インドネシア、ウルグアイ、オーストラリア、オーストリア、オランダ、カナダ、韓国、ギリシャ、スイス、スウェーデン、スペイン、スロバキア、セルビア・モンテネグロ、タイ、チェコ、チュニジア、デンマーク、トルコ、ドイツ、日本、ニュージーランド、ノルウェー、フィンランド、ブラジル、フランス、ハンガリー、ベルギー、ポーランド、ポルトガル、香港、マカオ、メキシコ、ラトビア、リヒテンシュタイン、ルクセンブルグ、ロシア (低所得国/参加国:0/41)	読解力、数学、科学、問題解決領域 15歳
TIMSS(1995年、1999年、2003年)	アメリカ、アルメニア、イギリス(イングランド)、イスラエル、イタリア、イラン、インドネシア、エジプト、エストニア、オーストラリア、オランダ、ガーナ、韓国、キプロス、サウジアラビア、シンガポール、スウェーデン、スコットランド、スロバキア、スロベニア、セルビア・モンテネグロ、台湾、チュニジア、チリ、ニュージーランド、日本、ノルウェー、パレスチナ、パーレーン、ハンガリー、フィリピン、ブルガリア、ベルギー(フラマン語圏)、ボツワナ、香港、マケドニア、マレーシア、南アフリカ、モルドバ、モロッコ、ヨルダン、ラトビア、リトアニア、ルーマニア、レバノン、ロシア (低所得国・地域/参加国・地域:2/46)	数学、科学 第4学年(小学校4年生相当)および第8学年(中学校2年生に相当)
PIRLS(2001年、2006年)	アイスランド、アメリカ、アルゼンチン、イギリス、イスラエル、イタリア、イラン、インドネシア、オランダ、オーストラリア、カタール、カナダ、キプロス、ギリシャ、クウェート、グルジア、コロンビア、シンガポール、スウェーデン、スペイン、スコットランド、スロバキア、スロベニア、ノルウェー、台湾、チェコ、デンマーク、ドイツ、トリニダード・トバゴ、トルコ、ニュージーランド、ハンガリー、フランス、ブルガリア、ペリース、ベルギー、ポーランド、香港、マケドニア、南アフリカ、モルドバ、モロッコ、ラトビア、リトアニア、ルクセンブルグ、ルーマニア、ロシア (低所得国/参加国:1/47)	読解(国語) 第4学年
SACMEQ(1998年、2000年)	ウガンダ、ケニア、ザンジバル・タンザニア、ザンビア、ジンバブエ、スワジランド、セイシェル、タンザニア、ナミビア、ボツワナ、マラウイ、南アフリカ、モウリシャス、モザンビーク、レソト (低所得国・地域/参加国・地域:8/15)	読解(英語)、数学 第6学年
PASEC(1995~2001年)	カメルーン、ギニア、コート・ジボアール、ジブチ、セネガル、中央アフリカ、トーゴ、ブルキナ・ファソ、マダガスカル (低所得国/参加国:8/9)	読解(仏語)、数学 第2学年、第5学年
LLECE(1997年、1999年)	アルゼンチン、キューバ、コスタリカ、コロンビア、チリ、ドミニカ共和国、パラグアイ、ブラジル、ベネズエラ、ペルー、ボリビア、ホンデュラス、メキシコ (低所得国/参加国:0/13)	言語(国語)、数学 第3学年、第4学年

(注)太字は低所得国。世界銀行『世界開発報告書』2006年版による定義で、購買力平価で一人当たり総国民収入(GNI)が2004年時点で825ドル以下を示す。

(出所)文部科学省ウェブサイト(http://www.mext.go.jp/b_menu/houdou/16/12/04121301/001.htm)およびhttp://www.mext.go.jp/b_menu/toukei/001/04120101.htm、SACMEQウェブサイト(<http://www.sacmeq.org/>)、World Bank(2006)IEAウェブサイト(<http://www.iea.nl/>)、World Bank(2006)、CONFEM(2003)、Michaelowa(2003)より筆者作成

表2 開発途上国の初等教育指標の現状

教育指標	地域	全世界	開発途上国	サブサハラ アフリカ	南西アジア
非就学児童数(百万人、2002/3年) (男子/女子)		99.9 (45.0/54.8)	95.5 (42.7/52.8)	30.1 (12.7/17.4)	40.4 (18.4/22.0)
純就学率(加重平均、2002/3年) (男子/女子)		84.6 (86.5/82.6)	83.2 (85.3/80.9)	82.5 (85.7/79.1)	63.5 (66.9/59.9)
留年率(全学年、%、2002/3年) (男子/女子)		3.0 (3.4/2.5)	6.2 (6.5/6.0)	4.9 (5.0/4.8)	15.5 (16.7/15.5)
退学率(全学年、%、2001/2年) (男子/女子)		15.1 (-/-)	25.5 (27.6/23.3)	35.1 (36.7/33.1)	40.5 (42.5/37.6)

(注)加重平均は、各国の学齢人口の規模に比例して各国のデータに重み付けをすることにより、地域別平均を算出したものである。人口が多い国ほど、地域平均への影響力は強くなる。

(出所) UNESCO(2006)のデータを基に筆者作成

質に関する政策的示唆を得ようとするようになる。南西アジアでは、その率は更に低く、全体の学齢児童のおよそ4分の1を母集団とした学力データに基づくことになる。そして、彼らの多くは社会の中でも比較的恵まれた子どもたちである。つまり、こうした定量データは中途退学した生徒や非就学児童を含む学齢児童全体の状況を反映したものではないため、一部の既に選ばれた子どもたちの状況を政策決定の基盤として用いることになる。

調査の入手方法としては、これら全ての調査は学校調査という方法を取っている。つまり、ある国の既存の学校がランダムに選ばれ、テストを実施した結果を用いている。ここには、退学している生徒や欠席しがちな生徒は含まれていない。

このように、国際的に行われる学力調査においては、低所得国の現状を鑑みた場合には、入手可能な教育評価の指標は限定されており、それらの政策的示唆にも一定のバイアスが発生することとなる。問題は、これらの定量的なデータが求められる政策にどこまで反映可能であるか、という点である。無論、現行の生徒に対する学校教育の現状を改善するという視点に立てば、学校に通っている生徒が、基礎科目において国が目指している基準をどの程度上回っているか、またはどのような学校あるいは生徒がどの程度の成績を取っているのか、を調べることに一定の政策的示唆を得る意義がある。しかし、学校に行っていない、あるいは中途退学した子どもたちにとっての学校教育の意味、あるいは彼らを救うための手立ての内容、に関する示唆を得るには十分ではない。このように、開発途上国における各国の教育政策を形作っていく上で、既存の国際的な調査があくまでも限られた方法による調査であることを認識しておくことが重要である。

3. 既存のデータ収集方法の有用性と限界：学校調査と世帯調査

前節では、既存の国際的な学力調査が、開発途上国に関して多くの情報を取りえていないこと、またその情報が多くの場合、一部の学齢児童のも

のに偏っていることを指摘した。これらは、学校調査という調査手法を取っていることも先に見たとおりである。本節では、教育評価において、学校調査と同程度使用される世帯調査の場合との比較において、現行の教育評価のための定量データの入手方法にいかなる課題が存在しているのかを考察する。

(1) 学校調査と世帯調査の目的と対象

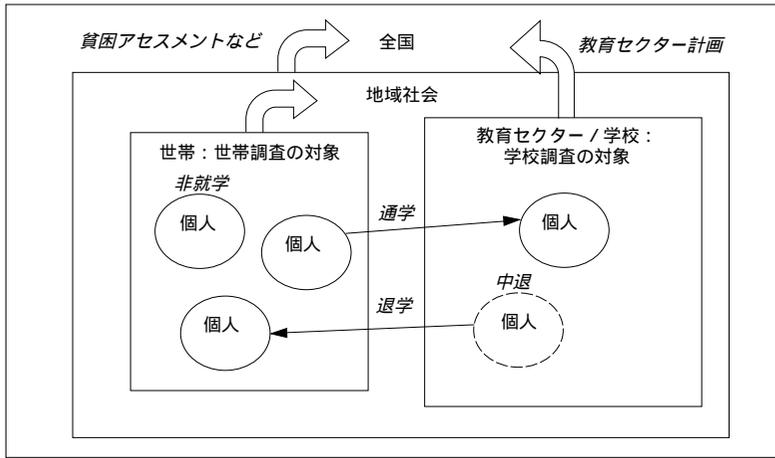
学校調査と世帯調査の目的は大きく異なっている。まず、学校調査は、国の全学校あるいはランダムに選ばれた学校から、生徒の就学状況、教員数、学校施設設備の状況、学校財政などに関する情報収集を通して、主に教育省などの政府が学校の置かれている状況を把握することを目的としている。国により、生徒の就学状況には、出席率、留年率、退学率などが含まれ、教員数も資格や学歴、性別ごとに情報が入手される場合もある。また、学校財政については、保有財産や授業料、生徒一人当たりが学校に対して支払っている授業料以外の学校経費などについて知ることが出来る。

他方、世帯調査は、世帯の支出収入を通じた富の度合いやそれらの使途や源泉、世帯主や世帯構成員が就いている職業、教育レベル、健康状態、家財や土地、家屋等の財産、貯蓄などの包括的な情報を基に、国全体の生活水準を把握するために、通常、政府が実施するものである。開発途上国、特に最貧国における世帯調査の殆どは、主に援助機関の支援を得て行われており、クラスター抽出法などを用いて確率標本により収集するため、全国民を代表するデータとされる場合が多い。

これら2つの調査の目的は、前者が学校教育に関するデータの入手を目的としているのに対し、後者は世帯構成員に基づく教育に対する態度や利用状況に関する情報入手を目的としているという意味で、教育に対する見方が異なっている。つまり、学校調査は教育の供給状態に焦点を当てているのに対し、世帯調査は教育の需要状態に焦点を当てている。

図1は学校調査と世帯調査の調査対象と目的を表したものである。世帯調査は世帯を単位としているため、世帯に含まれる子ども一人ひとりにつ

図1 学校調査と世帯調査の対象



(出所) 筆者作成

いての教育状況や世帯の環境を知ることができる。したがって、学齢児童であっても通学していない非就学児童についてもデータを入手することができる。また、パネルデータにすることにより、どのような子どもが留年しているか、あるいは退学しているかを経年的に観測することも可能である。但し、どのような学校に行っている子どもが退学や留年をしやすいのか、あるいはどのような出席状況や成績状況の子どもが退学や留年に至るのか、といった学校の情報については極めて限られた情報しか得られない。したがって、本人の属性や家庭の要因の影響が把握されたとしても、学校の状況を十分にコントロールしてないため、政策的示唆を得るには限界が生じる。つまり、ある子どもが学校へ行っていない場合、それは近隣の学校の質が悪いためか、家庭の事情によるものか、あるいはその複合的な理由によるものか、がはっきりしない。また、教育費などのデータも世帯を単位としているため、子ども別の教育費用を厳密に分けてデータ収集している世帯調査はまれである。このため、教育費用分析をする際に、世帯内のあるレベルの学校に通学する子どもの教育費を均一としたプロキシ（代替）指標の使用を余儀なくされる場合がある。

学校調査は、学校の成績や出席率などを一人ずつ観察することができる。また、一人当たり教育

費についても、生徒ごとに入手することが可能であり、世帯調査よりも厳密に算出することができる。したがって、学習の状況については、より細かいデータを有している。他方、生徒が退学した場合、退学率として取られるデータは、その特定の退学した子どもに関する情報ではなく、全体の生徒数に占める退学した子どもたちの割合（＝退学率）である。留年についても、当該年度のある学年の生徒の中で前年度も同学年にいた生徒の割合を以って留年率とすることが多い。したがって、収集される情報は、学校内の無数の生徒を母集団とした場合の留年や退学の傾向を知るものである。この場合、一人が何回留年しているか等、一人ひとりの行動の細部を追うことができず、家庭環境についても通常、殆ど情報は収集されない。さらに、ある生徒が退学したとしても、転校している場合や、別の学校に復学している場合などについての情報は厳密には追求できないこと、また母集団が新たに中途転入した子どもを含む場合もあるため、退学率や留年率についてのデータも厳密性を欠くことが多い。

上記の理由から、2つの調査方法から得られる政策的示唆は異なり、教育評価としての使用のされ方にも大きな違いがある。一般に教育評価に最も使用されるのは学校調査である。学校内で入手されたデータを以って教育セクター計画への示唆

を得るというものである。つまり、得られるべき「教育計画」はあくまで供給側から見た、学校整備、学校の質の改善である。そして先述したとおり、それらのデータは通学している生徒の環境やパフォーマンスを基に収集されたものを基準にしており、学校外の子どもたちは含まれていない。

他方、世帯調査は、教育セクター計画の参考とされることはあっても、その基盤になることはできない。なぜなら、その主要な目的は国民の生活水準を調査し、貧困の状況などを把握することであるからである。教育を享受すべき主要な公共サービスと捉えてはいるものの、学校に関する情報は殆ど取られていないため、教育セクターへの示唆を得ることは難しい。限られた活用法としては、世帯一人当たりの収入あるいは消費を算出し、各所得（支出）階層別の学齢児童の就学状況を比較することにより、奨学金やインセンティブの供与対象や提供方法の参考にすることがある。重要な点は、これらの調査のどちらも、教育へのアクセス、質、公平性を測るためには十分ではない、ということである。

(2) 得られる結果の相違

学校調査と世帯調査によるデータ収集方法の違いは、具体的に教育評価を実施するに当たり、その結果と教育政策や教育計画への反映という意味で、いかなる差を生むのであろうか。ここでは、ウガンダの教育スポーツ省が行っている学校調査（Uganda, MOES 2003; Uganda, MOES 2006）と筆者らが行ったウガンダ農村部の世帯調査に基づいた研究結果²を事例に考えてみたい。

具体例

まず、最も基礎的な学校へのアクセスのデータについては、2つの調査から得られたデータの間には若干の違いが見られた。学校調査で入手されたデータによれば、世帯調査により得られた就学率の方が、学校調査で得られた結果よりも若干低くなっている³。また、学校調査結果をよく見てみると、45県中女子で14県、男子で20県において純就学率が100%を上回っていた⁴。

また、学校調査結果によれば、2004年に全体の退学率は4.5%、留年率は13%となっている。教育スポーツ省は学年ごとの就学者数の構成から、1学年から2学年の間、および6学年から7学年が最も退学しやすい学年であるとしている（Uganda, MOES 2006）⁵。ところが、2003年および2005年に行った世帯調査のパネルデータによれば、これらの指標は学校調査により得られた結果よりも深刻な状況を示している⁶。また、退学と留年の傾向には男女間で大きな差が見られる⁷。

解釈

学校調査においては、就学率を算出する際に学齢人口を独自に調査するのではなく、人口推定値を使用している。このため、特に県ごとのデータともなると、人口移動や越境就学などの現状を細かく捉えることができないため、正確さを欠いた結果が得られるのだろう。世帯調査では、標本世帯における全人口を調査しているため、学齢児童については就学の有無にかかわらず、全員を調査の対象としている。ただし、世帯調査で仮に世帯主が子どもが学校に行っていると答えたとしても、厳密には学校での記録を確認する必要があるが、通常は行われていない。

また、留年率や退学率の差については、学校調査は学校を単位として集計しているのに対し、世帯調査は子どもを単位として集計していることに拠るものと考えられる。つまり、学校調査でいう留年率とは、前年度も同じ学年にいた生徒として、当該学年における留年者の割合を算出しており、前学年にいた生徒のうちの留年者の割合を算出しているのではない。また、留年者の定義として、前学年の途中で学校を退学し、次年度の途中から同学年に復学した生徒も対象としていることから、全体の母集団が一貫していない。これに対し、世帯調査においては生徒一人ひとりの動きを追っているため、母集団は一貫して、ある学年に就学していた生徒であり、彼らが2年間を経てどのように進学、留年、退学したのか、に注目している。この意味では、世帯調査の方がより厳密に生徒の進学状況を捉えることが出来るといえる。

次に、政策的示唆についても、世帯調査におい

ては、ある生徒の社会経済的な家庭環境を知りえているため、留年あるいは退学しやすい生徒の特徴を推定することが可能である。また、それを基に、どのような生徒を学校としてケアする必要があるのか、インセンティブや奨学金などの対象やその額の設定についても世帯の情報を基に推計することができる。ただし、学校の詳しいデータが不足するため、学校の学習環境に対する具体的な示唆を得ることは困難である。学校調査結果については、全国平均と比較した各学校の位置付けとそこからの対処指示は可能であるが、具体的に各学校でどのように対処したらよいのか、といった方策についての示唆を十分に得ることはできない。

このように、学校調査、世帯調査のどちらも異なる守備領域があるが、政策的示唆の観点からはどちらも教育政策や計画に反映できるほどの提言を出しえていないのが現状である。現状把握という点においては、どちらも国際目標に対する自国の位置を知り、課題を抽出することは可能である。しかし、状況改善という視点に立てば、国や地域に特有の改善策を生むだけの方向を提示しているとは言いがたい。

(3) 問題の所在

それでは、具体的に何が現状の教育開発の評価のデータ収集方法に欠けているのであろうか。表3は、先の議論を基に、教育へのアクセス、質、公平性という観点から教育評価における学校調査と世帯調査の有用性を比較し、それらに欠如している政策との関連に関する総合的教育評価モデルを提案するための基礎となるものである。まず、アクセスについては、学校調査と世帯調査ともに、就学率など大まかな状況を把握することは可能である。しかし、学校調査は子どもの社会経済的背景に関する情報、世帯調査は学校での学習環境に関する情報を欠いているため、就学状況にまつわる問題に関する家庭要因と学校要因を厳密に判断することができない。このため、非就学者が多いことが分かっても、具体的に取られるべき教育セクター内外における政策について明確な示唆が得られにくい。

教育の質については、世帯調査よりも学校調査の方がより詳細な情報を得ていると考えられるが、通常の学校調査は、教員数、教室数、などの教育の投入(input)に関するデータ収集に偏っており、結果(outcome)についての指標は殆ど収集されない。開発途上国の教育システムにおい

表3 教育評価における学校調査と世帯調査の有用性の比較と問題の所在

	学校調査	世帯調査	問題の所在
教育へのアクセス	登録者数、出席率などから確実に就学している生徒数を得ることができる。また、退学した生徒がどのような環境の学校に多いかなどの情報も有する。	就学していない児童を含む全世界構成員の教育へのアクセス状況と一人ひとりが置かれた家庭状況に関する情報を得ることができる。	就学していない児童に対する政策として、学校要因と家庭の要因のどこに問題があるのかを厳密に知ることができない。
教育の質	生徒の成績、学校の教員対生徒比率、教室当たりの生徒数、生徒一人当たりの教科書数等の学校の質についての情報を有し、学校単位での教育の質を知ることができる。また、退学者数や留年者数を学校単位で得ることができる。	学校教育の質については、公立私立、全日制、全寮制などの通学している学校のタイプや学校への距離などを知ることができ、家庭環境を踏まえた上で、それらと退学や留年との関連を推定することができる。	学校の質と退学や留年、教育の質と生徒の家庭環境の間の関連などについて明確な政策的示唆を得ることができない。
アクセスや質における公平性	性別、孤児などのグループごとに就学者数、退学者数、留年者数を知ることができる。	就学していない児童の属性や環境を、性別、孤児、親との同居状況、居住地域、親の職業や教育レベル、兄弟数などとの関連で知ることができる。	学校、家庭、教育の結果(outcome)に関する要因が社会のグループごとにどのように影響し合っているかを知るだけのデータが不足している。

(出所) 筆者作成

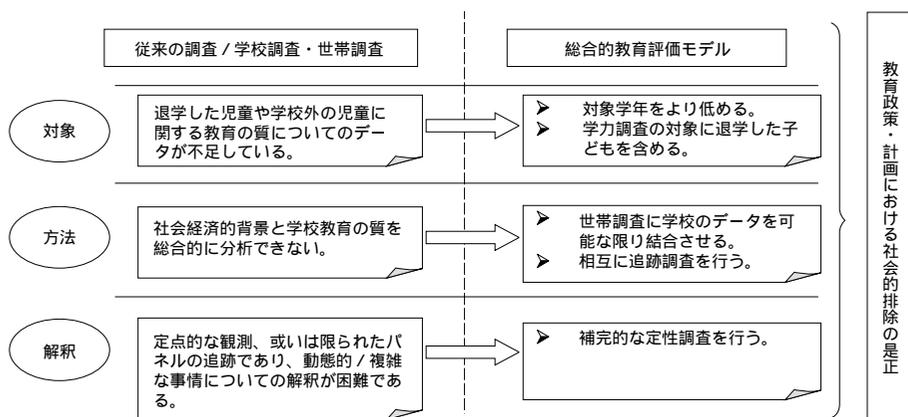
ては、初等教育修了試験を有する場合もあるが、それらの成績は最終学年の生徒についてのものであるだけでなく、生徒一人ひとりの学習状況や環境を把握するだけの、生徒を基準にした学習環境に関するデータは収集されない。このため、教育の質が投入の指標を以て測られやすいという欠点がある。先に見たとおり、成績といった教育の質の結果とされる指標はMDGsの対象とされていないことから、その評価方法への注目が比較的弱いこともある。Filmerらの研究によれば、修了率だけでは、初等教育で身に着けた能力を測ることができないだけでなく、実際の開発途上国の初等教育は社会に求められる能力を提供するに至っていないという（Filmer, Hasan and Pritchett 2006）。また、学校内だけでなく、退学者を含む学齢児童全体の能力を測定しなければ、実態に沿った教育の質を知ることができないことも指摘されている。したがって、この教育の質の測定方法については、学校調査、世帯調査の双方に改善の余地がある分野である。

公平性の観点からは、学校調査、世帯調査ともに男女、孤児、地域などのグループごとに教育のアクセスや質に関する情報を得ることはできる。しかし、学校調査と世帯調査が個別に実施されることが多いため、学校、家庭、教育の結果に関する要因について、グループごとにどのように連関しているか、を分析できるだけの体系的、包括的な情報は得られていない。

4. 結論 総合性の探求

前節で述べたとおり、教育評価に当たっては、教育指標に関するデータを収集するだけでなく、その要因となるような背景のデータを収集しなければ有効な政策や計画につながる評価を実施することは難しい。現行の国際的な学力調査や途上国内で実施されている学校調査や世帯調査は、対象や入手手法において各々限定されており、それ自体では具体的な施策に導くのは困難な評価調査である。重要な点は、先に述べた各調査が補えない部分をどのように結びつけて、全体像を推察するかである。その鍵となる視点は、図2に示すとおり、主に次の2点に要約される。第一に、教育評価に必要な視点に関する先進国と開発途上国の違いを認識することである。特に、学校外の子どもたちの多い低所得国においては、学校調査と世帯調査の結果に大きな乖離が生じやすいため、その双方を取り込んだ総合の必要性が大きい。また、教育セクター外の要因が大きく教育へのアクセスや公平性に影響している。このため、学力調査の場合、低所得層の状況を加味して学年をより低めるか、調査対象をすでに退学している、あるいは学校に行っていない学校外の子どもたちを可能な限り取り込む必要がある。また、世帯調査を実施する際に、各世帯の子どもが通う、或いは退学した学校名を聴き取り、学校の情報を結合させるのも一案である。

図2 総合的教育評価モデル



(出所) 筆者作成

第二に、開発途上国においては、様々な要因が教育へのアクセスや質に影響していること、留年や退学が一時点ではなく動的に起こることから、定性的な調査を含む限定的なサンプルを併用して、現状をより詳細に調査する必要がある。例えば、従来の定量調査の質問票に選択肢として掲載されている退学した理由などは、無関心、妊娠、結婚、授業料、仕事、病気、家事など、単純な理由ばかりである。このような選択肢は各国共通しており、集計結果は各国同じように見えてしまうが、より重要なことは、彼らが退学に至った過程や心境、学校に対する固有の考え方を知ることである。

一言に無関心、といっても、それが教育全般に関するものか、近所の学校に関するものか、あるいは十分に情報を持った状態で答えているのか、など明らかにされなければならない点が多い。また、仕事や家事といっても、地域的な特徴などの事情を把握しなければ、どのような教育が可能であるか、についての具体的な示唆を得ることができない。これらを聴き取るには、構造化インタビュー⁸を用いた定量調査だけでなく、半構造化インタビュー⁹などにより、定性調査を補うことが必要である。定性調査を並行的に行うことにより、地域的な特徴、個別の事情、貧困者の視点、学校教育に関する問題点などが洗い出され、より具体的かつ包括的な教育政策や計画に結びつく手がかりを掴むことができる。

定性的な手法を加えることの意味は他にもある。例えば、非就学者や退学者の脆弱性、不安定性や無力感といった心理的側面について知ることができる。また、世帯内の男女や年齢による資源の配分など家庭内の要素や、インフラや他の共有資源の欠乏などのコミュニティ全体で抱える要素をつかむことができる。さらに、就学の動態をより詳しく把握することができる。世帯調査は、学校調査とは異なりパネルデータ化できるが、それは毎年あるいは数年に一度の定点的な調査であり、食糧・雨量、健康状態、収入の季節的变化や長期的な傾向の影響や様々なショックに対する脆弱性をそれらと就学との関連の過程として観察することはできない。定性調査では、こうした変化に対して子どもたちが抱える問題のある程度補っ

て調査することができる¹⁰。

こうした認識枠組みとしての総合的教育評価を考案する目的は、画一的な教育計画の原因を探るだけでなく、社会的排除 (social exclusion) を行う教育政策や計画の策定を回避するためでもある。社会的排除とは、「市民的、政治的、社会的権利や機会へのアクセスの不足」(Buchanan 2006, p.1136)を指し、開発途上国においては、所得貧困や収奪 (deprivation)¹¹と同等ないしそれ以上に深刻な差別を意味している。中でも、教育が社会的排除に果たす役割は大きい (Meerman 2005)。教育計画の策定に当たって、学校に行っている子どもを対象とした学校調査のみを基盤とすることは、社会的排除の観点からも問題であり、包括的に社会構成員を対象とした世帯調査との総合性の視点が求められている。

暫定的な結論として、各国に合った教育計画を策定するためには、教育評価に当たっては、調査手法から得られる結果の総合性を重視し、開発途上国の事情を十分に加味した体制を整えなければならないと言う事が出来る。また、定量的なデータだけでは得られない推論 (inference) を定性的な調査から得ることも、非就学や退学などの問題に対するきめ細かい対応には必須である。Samoff (1999) や Klees (2001) がかつて教育セクタープログラムを批判していたように、国際的に画一的な教育政策・計画が作成されやすかった背景には、MDGsやEFAのような国際目標の共有によりもたらされた要素など幾つかの要因があるが、従来の各国政府における評価が非整合的に行われてきたという要素も存在しているのではないだろうか。

注記

- 1 UNESCO (2006). *Education for All Global Monitoring Report 2006*. Paris: UNESCOを参照。具体的な数値は表3を参照のこと。
- 2 筆者らは、2003年と2005年のウガンダの世帯調査により入手されたパネルデータを基に、初等教育へのアクセスと内部効率性に関する研究を行った (Nishimura, Yamano, and Sasaoka 2006; Nishimura, Yamano, and Sasaoka forthcoming)。
- 3 ウガンダの人口の約9割以上が農村部に居住してい

るが、この世帯調査は治安上問題のある北部地域および都市部を含んでいないため、完全な比較は困難である。

- 4 純就学率は学齢児童のうちで就学している生徒であるから、正確には100%を上回ることはない。
- 5 1学年が25%、2学年、3学年が其々16%、4学年が14%、5学年が13%、6学年が10%、7学年が6%となっている。
- 6 それぞれ退学率は13.4%、19.9%となっている。留年については、2年間で、34.8%の男子が一度留年し、30.5%が2度留年しており、女子では、それぞれ31.2%、31.4%となっている。
- 7 最も退学しやすい学年は、男子は初等最終学年の7年、女子は6年であり、最も留年しやすい学年は男女ともに2年生であるが、全学年を通じて留年率は高い。また、男女別の特徴としては、男子は6年生で留年した後、初等教育の卒業試験年次に当たる7学年で再度留年する率が高いのに対し、女子は一度6年生で留年した後は、留年するよりも退学する率が高い。すなわち、男子は留年しながらも卒業しようとする傾向があるのに対し、女子は留年を繰り返して一定の年齢に至ると退学のプレッシャーにさらされることが示唆されている。
- 8 得られた回答の間で比較ができ、作業が管理しやすく低労働ですむという利点を活かし、世帯調査などの定量的手法を用いた社会調査に広く使われるインタビュー法。高度に標準化された質問票を使い行われるもので、調査者の即興や自由度は殆どない。
- 9 構造化インタビューよりもインタビューする側の自由度を上げたインタビューの方法で、主に定性的な調査に使用される。
- 10 一般に、1990年代から定性調査の必要性が指摘されるようになり、貧困者の声 (Voices of the Poor) など参加型貧困アセスメントが実施されるようになったが、開発途上国の国家計画枠組みとしての貧困削減戦略ペーパー (PRSP) においては、未だにこうした定性調査の結果が十分に活かされていない。
- 11 収奪とは、各個人、家族、或いは集団が所属するコミュニティに比較して明らかに観察、説明しうる不利益のこと (Townsend 1993, p.79)。資源や一般的な環境ではなく、物理的、環境的、社会的な状態や状況としての特定の諸条件に当てはまることから、貧困とは区別される。これに対し、社会的排除とは

そもそも対象となる母集団にも入れない差別のことを指す傾向がある。

参考文献

- 西村幹子 (2006) 「ウガンダにおける初等教育政策の効果と課題：教育の公平性に注目して」『国際協力論集』、14 (2) : 93-117
- Brown, A., Foster, M., Norton, A. and Naschold, F. (2001). The Status of Sector Wide Approaches. *Working Paper no. 142*. London: Overseas Development Institute.
- Buchanan, A. (2006). Children Aged 0-13 at Risk of Social Exclusion: Impact of Government Policy in England and Wales. *Children and Youth Services Review, 28*: 1135-1151.
- CONFEMEN (2003). Elements d' appreciation de la qualite de l'enseignement primaire en Afrique francophone, background paper for the ADEA study on educational quality. Paris: ADEA.
- Filmer, D., Hasan, A. and Pritchett, L. (2006). A Millennium Learning Goal: Measuring Real Progress in Education. *Working Paper 97*. Washington, D. C.: Center for Global Development.
- Foster, M. (2000). New Approaches to Development Cooperation: What Can We Learn from Experience with Implementing Sector Wide Approaches? *Working Paper no. 140*. London: Overseas Development Institute.
- Heyneman, S. P. and Loxley, A. L. (1983). The Effect of Primary-School Quality on Academic Achievement across Twenty-nine High- and Low-Income Countries. *The American Journal of Sociology, 88 (6)* : 1162-1194.
- Klees, S. J. (2001). World Bank Development Policy: A SAP in SWAPs Clothing. *Current Issues in Comparative Education, 3 (2)*, 1-11.
- Meerman, J. (2005). Oppressed People: Economic Mobility of the Socially Excluded. *The Journal of Socio-Economics, 34*: 542-567.
- Levin, H. M. (1995). Raising Educational Productivity. In Carnoy, M. (Ed.). *The International Encyclopedia of Economics of Education*, 283-291. Pergamon.
- Levin, H. M. and McEwan, P. (2001). *Cost-Effectiveness*

- Analysis*, 2nd Edition. Sage Publications.
- Michaelowa, K. (2003). Determinants of Primary Education Quality: What can we learn from PASEC for francophone Sub-Saharan Africa?, background paper for the ADEA study: The Challenge of Learning: Improving the Quality of Basic Education in Sub-Saharan Africa. Paris: ADEA.
- Nishimura, M. Yamano, T. and Sasaoka, Y. (2006). Internal Efficiency under the Universal Primary Education Policy in Rural Uganda. Unpublished manuscript.
- Nishimura, M., Yamano, T. and Sasaoka, Y. (forthcoming, 2007). Impacts of Universal Primary Education Policy on Educational Attainment and Private Costs in Rural Uganda. *International Journal of Educational Development*, 27.
- Samoff, J. (1999). Education Sector Analysis in Africa: limited national control and even less national ownership. *International Journal of Educational Development*, 19 (4-5), 249-272.
- Townsend, P. (1993). *The International Analysis of Poverty*. New York, London, Toronto, Sydney, Tokyo, and Singapore: Harvester Wheatsheaf.
- Uganda, the Ministry of Education and Sports (MOES). (2003). *Technical Note on Primary Repetition, Survival, and Completion Rates before and after Universal Primary Education (UPE)*. in Uganda. Kampala: Ministry of Education and Sports.
- Uganda, the Ministry of Education and Sports (MOES). (2006). *The Education Management Information System (EMIS) data*. Kampala: Ministry of Education and Sports.
- UNESCO (2006). *Education for All Global Monitoring Report 2006*. Paris: UNESCO
- United Nations. (2004). *Implementation of the United Nations Millenium Declaration, Report of the Secretary-General*, No. A/59/282, United Nations.
- United Nations. (2006). *Millennium Development Goals Report 2006*. New York: United Nations.
- World Bank (2005). *Education Sector Strategy Update: Achieving Education For All, Broadening our Perspective, Maximizing our Effectiveness*. Washington, D.C.: World Bank.
- World Bank (2006). *World Development Report 2006: Equity and Development*. Washington, D.C.: World Bank.

(2007.2.14受理)

A Theoretical Comparative Analysis on Educational Evaluation in Developing Countries: Perspectives of International Student Assessment, School Survey, and Household Survey

Mikiko Nishimura

Graduate School of International Cooperation Studies, Kobe University
mnishimu@kobe-u.ac.jp

Abstract

The article attempted a theoretical comparative analysis on educational evaluation methodologies to examine the ways in which educational evaluation can lead to effective education policies and planning in developing countries. Three major surveys, namely, international student assessment, school survey, and household survey were compared and contrasted and the missing links were identified. It was found that the existing international student assessments have defects in its target groups and data collection method, while school survey and household survey also overlook a comprehensive focus on both supply and demand sides of education. The results of these surveys occasionally confront discrepancies, especially in developing countries where a number of out-of-school children and dropouts exist. Thus, a comprehensive assessment of these surveys is essential. Furthermore, qualitative research would be required for a complementary analysis for inferences from the quantitative data, particularly for critical issues such as non-enrollment and dropout.

Keywords

educational evaluation, developing countries,
quantitative data, school survey, household survey

【研究ノート：依頼原稿】

教育分野における参加型開発支援プロジェクトの 評価に関する基礎的考察

石田 洋子

財団法人国際開発センター

ishida.y@idcj.or.jp

要 約

途上国では、限られた資源で初等教育改善を進めるために、住民参加による学校建設や学校運営改善などの活動が展開されている。こうしたコミュニティの努力を、就学率向上につなげるには、住民の活動に対し、途上国政府が、教員配置や教科書配布、学校施設維持管理への技術指導などの行政サービスを提供して、教育の質を確保することが不可欠である。日本の政府開発援助（ODA）による教育分野の技術協力では、住民参加による教育改善活動をパイロット・プロジェクトとして行い、地方分権政策下で必要な現地教育行政機関の組織強化や能力向上が図られている。これらのプロジェクトは、行政官の教育行政能力向上に重点を置き、教育開発の視点から評価されているが、住民の参加型開発を支援する行政官の能力向上にも効果があることから、参加型開発の視点からも評価され、フィードバックを参加型開発推進のための環境整備につなげることが重要と考えられる。

キーワード

住民参加による初等教育改善、参加型開発への行政支援、参加型開発のための環境整備

1. はじめに

2015年までにミレニアム開発目標（MDGs）を達成するという国際社会の合意に基づいて、教育分野では、「普遍的初等教育」を目指して、途上国およびドナー（国際機関および二国間援助機関）による努力が展開されている。しかし、2005年の中間報告によると、教育分野は、他分野に比べてMDGs達成の進捗が遅く、2015年までに初等教育の就学率100%が達成される見込みは低いとされる。特に、サハラ以南のアフリカおよび南西アジアにおいて、その達成が危ぶまれている（国際協

力機構 2005b）。

こうした背景のもと、多くの途上国において、限られた資源を最大限に利用して、初等教育へのアクセスを拡充するために、住民参加による学校建設などの初等教育改善プロジェクトが展開されている。

日本の政府開発援助（ODA）でも、住民参加を導入し、強化することによって、初等教育の量と質の改善を目指す技術協力プロジェクトが実施されている（国際協力事業団 2002a）。これらのプロジェクトは、二国間援助の特性から、現地の教育行政機関および行政官を直接の技術移転の対

象グループとし、現地行政のデータ管理や計画策定における人材育成や組織強化をプロジェクト目標とし、教育開発が進むこと、教育指標が改善することを上位目標とする（国際協力事業団 2002b、2002c、JICA 2002）。したがって、プロジェクト評価の際には、教育開発の視点から評価が行われる。

しかし、これらのプロジェクトの多くは、JICA専門家やコンサルタントが、現地行政官と協働して、住民に直接に働きかけるパイロット・プロジェクトを実証調査として実施し、現地行政官が参加型開発の指導や支援を実践する機会を提供しており、参加型開発の上でも重要な役割を果たしていると考えられる。パイロット・プロジェクトでは、住民は、計画作成や学校建設サイトを決める意思決定の段階、実際の活動を行う実施段階、ならびにモニタリング・評価段階に参加する（国際協力機構 2004）。住民参加による持続的な開発の体制づくりを支援しながら、教育の質を確保するには、教育行政が、住民による教育改善活動の計画段階、実施段階、そしてモニタリング・評価段階において、それぞれ情報提供や技術指導を行うことが必要である。また、学校建設などのプロジェクト終了後も、教員配置、教科書配布、学校運営指導など、教育行政の関与は不可欠と考えられる。

住民による主体的な参加による教育開発と、教育行政をどのように結びつけるか、教育行政の関与や支援は、どのタイミングで、どのような形態で行われるべきか。実証調査を含む教育分野の技術協力プロジェクトは、現地行政が参加型開発支援のあり方を経験し、問題点やポテンシャルなどを具体的に検討する貴重な機会を、途上国行政機関にも、日本側にも提供しているといえよう。こうした機会を活用するためには、教育開発の視点からだけでなく、参加型開発の視点からも、当初から、目標や指標を設定し、モニタリングと評価が行われることが必要と考える。

住民の主体的な参加型開発が推進されるためには、住民による意思決定が尊重される民主的な政治体制が中央や地方行政レベルで保障されていることや、住民が地域の情報や資源に自由にアクセスでき活用できるよう、行政組織や制度面での環

境整備も必要となる（斉藤 2003、佐藤 1997）。技術協力プロジェクトの評価を通して、改善のポイントや阻害要因を明らかにすれば、現地行政のグッド・ガバナンスにつなげることも可能となる。また、具体的なフィードバックによる行政の改善は、教育分野でドナー協調により進められている直接財政支援の途上国側の実施体制を拡充する上でも有益であると考えられる。

本稿では、住民参加による教育プロジェクトの現状を概観し、参加型開発支援のパイロット・プロジェクトを含む、教育分野の技術協力プロジェクトの実施状況を示した上で、これらの参加型開発支援プロジェクトに対する評価のあり方を検討し、マラウイにおいて実施された技術協力プロジェクト（開発調査）の事例研究を通して、提言を取りまとめた。

2. 住民参加による教育プロジェクトの現状

多くの途上国では、学校数、教室数が圧倒的に不足しているにもかかわらず、教育予算は、その大部分が、行政官や教員の給与を中心とするリカレント予算に費やされる（国際協力事業団 2001、2002c、国際協力機構 2005c、JICA 2002、2005、2006）。このため、開発予算がごくわずかであり、学校建設をカバーできない。一方、ドナー支援は、必ずしも支援ニーズが高いところに配分されるとは限らず、現地行政の政治的な思惑や、支援側にとってアクセスの比較的良好なところなどに、資源が集中して配分される傾向がいまだに強いと思われる（外務省 2005）。

こうした状況においてアクセス改善を進めるため、途上国政府の中には、ほとんど現地行政が関与しないで、住民参加によってコミュニティ・スクール建設や教室建設を進める政策をとっているケースがある（国際協力機構 2005a）。例えば、エチオピアでは、就学人口に対して小学校数が極端に少なかったことから、このギャップを埋めるため、計画作りから、資金や資機材調達、建設作業、教員の配置まで、全て住民参加によって初等学校の建設・維持管理を行う、代替基礎教育センター（Alternative Basic Education Center）の設置

を推進する政策が実施されている（国際協力事業団 2003a、JICA 2006）。シエラレオネでは、紛争で破壊された教育施設を建て直すために、住民参加によって、フォーマル校を代替するコミュニティ・スクール建設を進めることが政策として掲げられている（国際協力機構ホームページ）。

これらの学校の多くは、建設基準とは関係なく、政府による工程監理や技術指導も行われないうちに、地域の労働者を雇って建設されることから、施設の耐久性や安全性に問題が多いとされている。さらに、コミュニティ・スクールには、正規の教員が政府から配属されないケースが多く、地域の比較的教育レベルの高い者が指導者に選ばれ、授業を受け持っている。給与は、コミュニティからのわずかな寄付金が農作物が当てられる。多くの場合、教科書や教材の配布も十分ではない（JICA 2005、2006）。

地域の未就学児童の状況を踏まえて、学校建設の計画を作成したり、計画段階で、教員配置や教科書配布をどうするか検討したり、建設作業や維持管理の段階においても、教育行政からの関与は非常に限られていると考えられる。住民参加による学校建設というと住民の結束があるようで響きが良いが、行政サービスを住民に肩代わりさせるだけの住民参加ともいえよう。住民任せの学校建設では、一時的に就学者数を増やすことには役立つかもしれない。しかし、そこで提供される教育の質は低く、やがて、中退者の増加や修了率の低下が顕在化し、長期的改善にはつながらないと考えられる（JICA 2005、2006）。

ドナー支援や政府資金が一部提供されて学校建設が行われる場合にも、労働が資金など何らかの形で住民負担が3割以上あることなど、住民参加が資金援助の必要条件とされることもある。例えば、マラウイでは、住民グループが、学校建設のプロポーザルを提出すれば、建設資金援助が得られるが、住民負担が一定以上あることが必要とされる基金が政府によって運営されている（世銀援助）。資金援助はあるものの、計画作成や、建設、維持管理や学校運営の段階での、現地行政の関与や支援は、ここでも非常に限られており、住民任せの状態にあると考えられる（国際協力機構 2005c、JICA 2005）。

NGOによる住民参加の学校建設プロジェクトの場合は、上記のような現地行政があまり関与しない住民参加の学校建設プロジェクトに比べると好対照で、計画作成段階から、技術支援が行われることが多い。このような場合は、教育開発における住民参加によって、学校が建設され、アクセスが改善されるだけでなく、住民のエンパワーメントにつながり、コミュニティ開発への波及効果も生みだすことが報告されている（CanDo 2004、JNNE 2004、TCSF 2006）。

途上国におけるドナーやNGO支援による学校建設プロジェクトでの経験に加え、日本における明治以降の教育開発の経験に基づいて、教育開発を、住民参加で行うことの主な効果として、以下が挙げられる（国際協力事業団 2003b、TCSF 2006）。

- (1) 住民が学校建設や資機材調達に参加することによって、教室建設の費用が節約でき、限られた資金で、より多くの学校・教室が建設できる。
- (2) 住民が教室建設や維持管理に加わることによって、学校へのオーナーシップや、教育の重要性への理解が高まり、子供達を学校へ送るようになる。
- (3) 住民が学校運営に加わることによって、学校で行われる授業や活動内容やレベルに関心が高まり、教育の質の向上につながる。
- (4) 住民による主体的な参加により、教育分野での開発協力を実施することによって、地域住民のエンパワーメントにつながる。

途上国の政府側からは、住民参加を進めるメリットとして、(1)が強調されるケースが多い。しかし、本来の住民参加の目的は、コスト削減だけではなく、住民のオーナーシップや教育への理解を高めるためであり、住民のみで(2)、(3)、(4)につなげることは、NGOによる支援が入ったケースを除くと、現状では難しい状況にある。

参加型開発の対象を初等教育改善とすることは、農業開発や生計向上プロジェクトに比べて、以下のような利点があり、参加型開発を手段として、住民のエンパワーメントを進めるエントリーポイントとして、コミュニティで受け入れられやすいと考える（TCSF 2006）。

- (1) 経済活動を目的としないことから、特定の個人やグループの利益につながりにくい。
- (2) 地域の子供の教育環境を良くするという共通の、合意しやすい目的を持つ。
- (3) 新しい技術や機材を必要とせず、住民の体力にあったプロジェクトが考えられる。
- (4) プロジェクトで子供の教育環境がよくなり、分かりやすい成果が上げられる。
- (5) 地方教育事務所、学校、住民のスクール・コミッティなど既存の組織が存在する。

3. ODAによる参加型開発支援の意義

NGOの支援活動には限りがあり、途上国の全てをカバーできるわけではない。そこで、住民の主体的な参加型開発活動に支援を行う役割は、地方レベルの現地行政が担うことが期待される。こうした背景を踏まえて、住民の声を反映した教育計画づくりや、住民参加による学校建設への行政による技術支援、学校運営のための行政指導や住民参加の効果的モデル作りを目的とした技術協力

プロジェクトが、JICAによって実施されている（国際協力事業団 2004）。本稿では、このような教育分野における技術協力プロジェクトを、「教育参加型開発支援プロジェクト」と総称する。

教育参加型開発支援プロジェクトには、インドネシアにおける「インドネシア地域教育開発支援調査（REDIP）」¹⁾、マラウイ「全国地方教育支援計画策定調査（NIPDEP）」²⁾、エチオピア「住民参加型基礎養育改善プロジェクト（ManaBU）」³⁾、ニジェール「住民参画型学校運営改善計画（みんなの学校プロジェクト、Ecole pour Tous）」⁴⁾などの技術協力プロジェクトが含まれると考える（国際協力事業団 2003a、国際協力機構 2006、JICA 2001a、2005）。

NGO支援によって、参加型開発プロジェクトを進めることによって、教育改善だけでなくコミュニティ・エンパワーメントの成果が上がっていることが報告されている。では、ODAによる技術協力プロジェクトによって、参加型開発支援を行う意義は何であるのか。

表1に、ODA支援による参加型開発支援プロジェクトの目的、対象グループ、投入規模などの傾

表1 ODAとNGO支援による教育参加型開発支援プロジェクトの傾向比較

	ODAによる 参加型開発支援プロジェクト	NGOによる 参加型開発支援プロジェクト
プロジェクト 目標	・教育の量と質の改善 ・教育行政機関の組織強化と、教育行政官の能力向上	・教育の量と質の改善 ・コミュニティのエンパワーメント
協力提供者	・JICA 専門家またはJICAから委託を受けたコンサルタント・チーム	・国際または現地NGOスタッフ
カウンターパート (担い手)	・中央または地方教育行政	・住民グループ
主要な活動	・中央・地方政府組織や行政官に対するキャパシティ・ディベロップメント（研修と実践） ・パイロット・プロジェクトの実施（学校建設、現職教員研修、教科書配布、学校運営改善など）	・対象グループとの学校建設や学校運営改善プロジェクトの計画作成と実施
対象グループ	・地方教育行政官 ・パイロット地域の学校関係者、スクール・コミッティのメンバー、住民	・対象コミュニティの学校関係者、スクール・コミッティのメンバー、住民
投入	・JICA 専門家やコンサルタント・チームの人件費 ・研修やプロジェクト運営費用 ・パイロット・プロジェクト費用	・NGOスタッフの人件費 ・学校建設などにかかる活動経費

(出所)筆者作成

向を、NGOの参加型開発支援プロジェクトのそれらと比較した。

両者とも、住民参加を取り入れて、教育の量と質の改善を目指すとともに、コミュニティのエンパワーメントも視野に入れた活動を行う点では同じであるが、ODAプロジェクトでは、プロジェクト目標には、教育行政官の能力向上や、教育の質や量の改善など、教育開発に重点がおかれることが多く、NGO支援プロジェクトでは、住民のエンパワーメントもプロジェクト目標に含まれると考えられる。

ODAプロジェクトを実施するJICA専門家やコンサルタントと、NGO支援プロジェクトを実施するNGOスタッフは、いずれも、対象コミュニティに対して外部者の立場をとることは同じであるが、受益者である住民に対する位置づけは異なる(図1)。

NGO支援プロジェクトでは、プロジェクトのカウンターパート(担い手)は、住民グループであり、外部者であるNGOスタッフが、住民グループに働きかけて、参加型開発の活動を支援する。

一方、ODAによる教育参加型開発支援プロジェクトでは、住民に対する参加型開発の技術支援を、地方教育行政官による教育行政サービスの一環と位置づける。外部者である専門家やコンサル

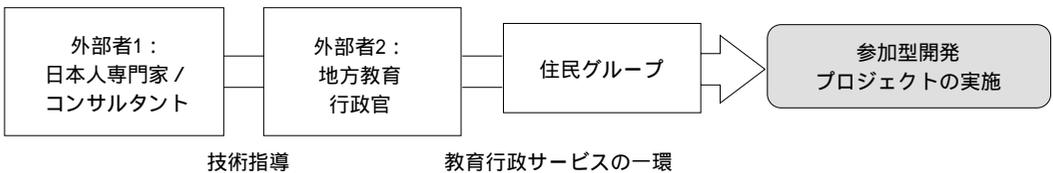
タント・チームは、より住民に近い立場の外部者である地方教育行政官に技術支援を行い、地方教育行政官が、住民グループに対して働きかけを行う(JICA 2001、2005)。

NGO支援プロジェクトも、ODA支援による教育参加型開発支援プロジェクトも、特定のコミュニティに働きかけて、住民参加による教育開発プロジェクトを支援するという点は同じである。ODA支援プロジェクトでは、JICA専門家やコンサルタントは、カウンターパートである中央または地方行政機関と協働する。そして、地方教育行政官を通して、住民に働きかける。このため、中央及び地方教育行政官の参加型開発に対する理解を深めるとともに、自信を強め、住民主体の参加型開発を支援する行政サービスを拡充する機会を持つといえよう(国際協力機構 2005、JICA 2005)。

ODAによる教育参加型開発支援プロジェクトには、パイロット・プロジェクトのプロジェクト・サイクルと、行政官のキャパシティ・ディベロップメントを行う本体プロジェクトのプロジェクト・サイクルの2つが含まれ(図2)、前者は、後者の「実施とモニタリング」の段階に組み込まれると考えられる。パイロット・プロジェクトからは、参加型開発に対する行政支援のあり方を明らかにする上で、貴重なフィードバックを得ることが可能である。

図1 ODA支援とNGO支援による外部者の住民グループへの位置づけ

【ODA支援のケース】

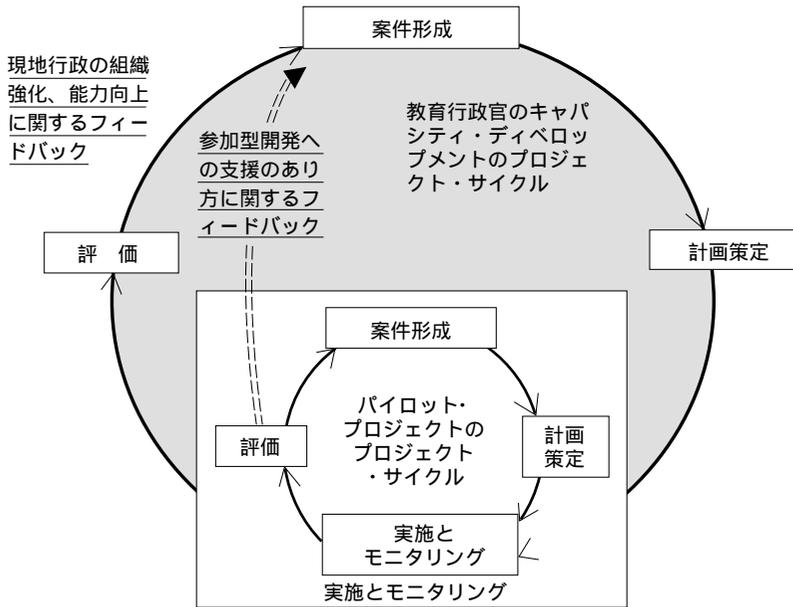


【NGO支援のケース】



(出所) 筆者作成

図2 教育参加型開発支援プロジェクトの2つのプロジェクト・サイクル



(出所) 筆者作成

こうした教育参加型開発支援プロジェクトの意義、プロジェクトの構成などを踏まえて、プロジェクト目標を設定し、参加型開発の視点からもモニタリング・評価が行われるように評価デザインを設定して、適切なフィードバックを得なければ、教育分野において、ODAの技術協力で参加型開発支援プロジェクトを展開する意義が損なわれるといえよう。

4. 参加型開発への支援強化に対する評価の重要性

教育参加型開発支援プロジェクトの多くは、教育行政官の能力を向上させる機能に加えて、参加型開発の環境整備を行う機能を有し、既に成果も上げていると考えられる。しかし、ODAの「教育分野の技術協力」の枠組みの中で実施されていることから、教育行政官の能力向上という前者の機能は重視されるが、参加型開発支援という後者の機能に関連する目的や活動は、プロジェクト・

デザインに記載されず、専門家やコンサルタントからも意識されないケースが多いと思われる。

これらのプロジェクトの評価では、教育行政官に対する能力向上や、教育行政機関の組織強化の達成度が、プロジェクト目標の指標として設定され、評価される。上位目標には、教育プロジェクトの実施状況や教育の量や質の改善状況が設定され、就学率や中退率などの教育指標改善に対する貢献度について評価されるケースが大部分であると考えられる。

しかし、参加型開発のアプローチをとり、パイロット・プロジェクトによる実証を活動に含めるのであれば、参加型開発に対する行政の役割や行政の支援機能に関する指標を取り入れて、モニタリング・評価し、参加型開発に対する行政支援のあり方や、参加型開発を進める上の阻害要因などをフィードバックとして得ることによって環境整備につなげることが重要であろう。

表2 マラウイ全国地方教育支援計画策定調査の概要

プロジェクトの要約	
長期上位目標	マラウイの初等・中等教育の質・量両面での改善が進む
中期上位目標	地域のニーズを反映して作成・更新された県別教育開発計画（DEP）に基づいて初等・中等教育の改善が進められる
プロジェクト目標	地方教育行政官のデータ管理、DEP作成、実施、モニタリング・評価能力が向上する
主な活動	地方教育行政官に対するDEP作成やデータ管理能力向上のための研修ワークショップの実施 DEP実証調査として、パイロット県における、パイロット・プロジェクトの実施 DEPを活用し、中央政府が、地方教育開発活動を支援するための「全国地方教育支援計画」の策定

（出所）国際協力機構 2005、JICA 2005

5. 事例研究：マラウイ地方教育支援計画策定調査に対する評価

本セクションでは、2003年1月から2005年10月まで、JICAの技術協力プロジェクトとして実施された「マラウイ全国地方教育支援計画策定調査（開発調査）」を事例として取り上げ、同プロジェクトの最終報告書に基づいて、参加型開発および参加型開発への行政支援の視点からの評価を試みる。

（1）対象プロジェクト「マラウイ全国地方教育支援計画策定調査」の概要

マラウイでは、初等教育への量的拡大を目指して、1994年に初等教育無償化政策が導入された。総就学率は急速に増大したものの、教室や教員、教材の不足などから、中退率の増加や修了率の低下など、教育の質の低下が深刻となった。マラウイ教育省は、限られたリソースを生かして、貧困削減やMDGs達成につなげるためには、地方分権化政策に沿って、地方教育行政官のデータ管理、計画作成、実施、モニタリング・評価の能力を向上すると共に、地方教育行政官の監督の下、住民参加によって、教室建設や学校運営改善を推進す

ることが不可欠と考え、JICAに技術協力を要請した（国際協力事業団 2002、国際協力機構 2005c、JICA 2002、2005）。

こうした背景から、2003年2月より2005年9月まで「マラウイ全国地方教育支援計画策定調査（開発調査）」が実施された¹⁾。同プロジェクトの要約は表2の通り。

（2）参加型開発支援機能に関する評価

対象プロジェクトでは、中央及び地方教育行政官に対する計画作成、実施、モニタリング、評価などの研修やOJTが行われると共に、地方教育行政官のモニタリング、技術指導のもとに、パイロット県においてパイロット・プロジェクトが実証調査として実施された。

パイロット・プロジェクトでは、住民参加が奨励された。地方教育行政官が技術指導を行い、パイロット・プロジェクトごとに結成されたタスクフォース（学校関係者や住民グループ）が、パイロットの運営管理、財務管理を行った（国際協力機構 2005c、JICA 2005）。

これらの成果を踏まえて、地方教育開発活動を支援するための制度整備の一環として、「全国地

方教育支援計画」が策定され、マラウイ教育省に提出された。その目標や活動内容から、対象プロジェクトは、教育参加型開発支援プロジェクトであるとともに、制度整備の重要性を意識したプロジェクトであったといえよう。

対象プロジェクトでは、地方教育行政官によるパイロット・プロジェクトのインパクト調査、パイロット・プロジェクト終了時評価、キャパシティ・ディベロップメントに関する終了時評価と、3つの内部評価が実際に行われた（国際協力機構 2005c、JICA 2005）。

地方教育行政官によるインパクト評価結果には、参加型開発を行うことの有効性や制約などが記載されている。しかし、日本人コンサルタント・チームが、現地コンサルタントに委託して行ったとの評価には、パイロット県の教育データの変化や、行政官の計画作成や実施能力の自己評価などが中心であり、参加型開発アプローチを導入したことによる変化や、そこから得られた教訓などは記されていない。これは、対象プロジェクトの目標が、PDMにも示されている通り、教育行政官の能力向上であり、上位目標が教育の質と量の改善であったためと考えられる。

対象プロジェクトでは、2年次に亘り、教室建設、トイレ建設、教員宿舎建設、井戸建設などの学校施設整備、現職教員再研修、住民啓発キャンペーン、教科書配布など、80件近くのパイロット・プロジェクトが、パイロット県6県で実施された（国際協力機構 2005c、JICA 2005）。

パイロット・プロジェクトは、地方分権化の中で、地方教育行政官の能力向上の場とすることが優先されたため、参加型開発を導入することがどのように検討され、合意されたかは、報告書には特に示されていない。また、どの段階での住民参加が、教育改善に対してより効果的かについても明らかではない。これらの点が明らかになるような形で実証調査が計画されれば、フィードバックはより役立つものとなると考えられる。

住民参加は、トップダウンでプロジェクトを進めるよりも時間がかかること、一般的に、途上国では、人件費や現地調達可能な建設材料費などは高額でないことから、住民参加が必ずしもコスト削減につながらないことなどが、報告書には示

されている。同時に、住民参加により、住民の自信が高まり、住民が、独力で新たな学校施設建設や井戸建設を行ったり、自主的にマッチング・ファンドを集めてより多くの教科書を購入できるように工夫したりと、自主性やオーナーシップの高まりが見られた（国際協力機構 2005c、JICA 2005）。長期的にみて、こうした成果を生み出し、育てることは、参加型開発導入時の短期的な効率性の悪さを、十分にカバーするものと考えられる。

マラウイ全国33教育県のうち、パイロット県6県でパイロット・プロジェクトが実施された。33教育県は6つのディビジョンに分けられており、1つのディビジョンから1つのパイロット県が選定された。地方教育行政官のOJTとしてみた場合、今後、パイロット県以外で、マラウイ政府によって同様のOJTが行われることは難しく、より多くの県がカバーされることが期待される。

地方教育行政官によって実施されたインパクト評価の結果には、参加型開発をとることによって、住民の意欲が高まったこと、教育改善に効果があったことが盛り込まれており、参加型開発に対する理解は高まったと考えられる。しかし、地方教育行政官自らが、どのような役割を果たす必要があるのかについては、まだ明確に認識されていないといえよう。

対象プロジェクトでは、「全国地方教育支援計画」を作成し、住民ニーズや教育データを反映させて地域レベルの教育開発計画を作成・更新し、実施する恒常的しくみをマラウイ教育省に示した。こうした計画に基づいて、地方レベルの参加型開発に対する支援体制をより強化し、参加型開発の阻害要因を排除していくことは、住民の主体的参加による持続的開発の体制を築くには不可欠と考えるが、さまざまなレベルでの意識改革や権限委譲などが必要になると考えられ、外部からの長期的な支援が望まれよう。

6. 参加型開発の視点を評価にもりこむための提言

マラウイにおける教育参加型開発支援プロジェクトにみられるように、教育分野の技術協力プロ

プロジェクトには、参加型開発を支援し、コミュニティのエンパワーメントを推進するには、途上国政府側に、組織強化、制度整備など何が必要なのかを、現場での経験や実証を通して明らかにできる有益な機能が含まれている。

マラウイの事例研究に基づいて、教育参加型開発支援プロジェクトを、参加型開発の観点から評価する際には、表3に示す評価の視点をを用いて分析し、貢献要因や阻害要因を抽出し、教訓を取り

まとめることで、効果的なフィードバックが得られると考えられる。

上記のような視点を、教育開発の視点に加えて評価する必要性を念頭において、プロジェクト目標や指標を設定し、事前評価、中間評価、終了時評価、そして事後評価が行われることが望まれる。教育参加型開発支援プロジェクトから得られたフィードバックを通して、参加型開発が進められ、市民社会が強化されれば、グッドガバナンスや民

表3 参加型開発に関する評価の視点

	キャパシティ・ディベロップメント に対する評価	パイロット・プロジェクト に対する評価
妥当性	<ul style="list-style-type: none"> ・教育開発に、政府の政策として、参加型開発を導入し、推進する妥当性はあったか ・参加型開発を導入することに、相手国政府との合意はあったか ・本体プロジェクトに対し、実証調査の目的、規模、内容、スケジュールは妥当であったか 	<ul style="list-style-type: none"> ・住民参加で教育開発プロジェクトを実施することは、住民のニーズや価値観からみて妥当性はあるか ・参加型開発の導入は、既存の組織体制などからみて、妥当性はあるか
有効性	<ul style="list-style-type: none"> ・パイロットを通して、地方教育行政官の参加型開発への理解と、参加型開発を支援する自分の役割の認識は向上したか ・地方教育行政官による参加型開発への支援能力（ファシリテーション、技術指導など）は向上したか 	<ul style="list-style-type: none"> ・意思決定の段階から、住民参加が組み込まれたか ・案件形成、計画、実施、評価のどの段階で、どのようなタイミングで、どのような形で住民参加を入れることが有効であったか ・住民の参加は、自由意志によるものだったか、強制参加であったか
効率性	<ul style="list-style-type: none"> ・JICA専門家やコンサルタント・チームの投入量やタイミング、専門性や能力は、参加型開発の支援強化を進める上で効果的であったか 	<ul style="list-style-type: none"> ・住民参加を入れることで、時間的効率性が変わったか ・住民参加を入れることで、費用的効率性が変わったか
インパクト	<ul style="list-style-type: none"> ・地方教育行政官の参加型開発に対する意識や行動様式は変わったか ・地方教育行政官は、自主的に、参加型開発への支援活動やフォローアップ活動を行うようになったか ・地方教育行政の行政能力や透明性は改善されたか 	<ul style="list-style-type: none"> ・住民の学校改善や教育開発プロジェクトに対するオーナーシップは強化されたか ・コミュニティ開発に対する住民の意識や行動様式はどう変わったか ・住民と現地行政の関係は強化されたか
自立発展性	<ul style="list-style-type: none"> ・中央行政、地方行政の両方のレベルにおいて、参加型開発に対する理解と支援が進んでいるか ・参加型開発を阻害する制度や体制について改善が検討されているか ・参加型開発を支援・拡大する制度や体制、または基金などのしくみが計画または構築されているか 	<ul style="list-style-type: none"> ・コミュニティ開発のために住民グループなど組織体制が生まれたか ・教育プロジェクトを契機に、新たなプロジェクトが計画、実施されたか

(出所) 筆者作成

主化の実現にもつながるであろう。そうなれば、教育分野のみならず、セクターを越えた参加型開発支援プロジェクトに対して、参加型開発の視点からの評価を徹底し、フィードバックを得ることの重要性が、ますます高まるものと考えられる。

注記

- 1 同開発調査に先立ち、2000年11月～2002年7月に、「マラウイ全国スクールマッピング・マイクロプランニング（開発調査）」が実施され、地方教育行政官のデータ管理・計画作成能力を向上し、県ごとの教育開発計画を作成する技術協力プロジェクトが実施された。

参考文献

- アフリカ地域開発市民の会（CanDo）（2004）『CanDo アフリカ地域開発市民の会会報1998年2月～2003年12月』アフリカ地域開発市民の会
- 教育協力NGOネットワーク（JNNE）研究会報告書（2004）『どう終えた？どう終える？プロジェクト撤退のポイント』外務省経済協力局民間援助支援室
- 外務省（2005）『教育関連MDGs達成に向けた日本の取り組みの評価』（評価報告書）外務省経済協力局
- 国際協力事業団（2001）『インドネシア国地域教育開発支援調査 最終報告書 - 和文要約』国際協力事業団
- 国際協力事業団（2002a）『開発課題に対する効果的アプローチ - 基礎教育 - 』国際協力事業団
- 国際協力事業団（2002b）『タンザニア地方教育行政強化調査（スクールマッピング・マイクロプランニング）ファイナル・レポート 要約』国際協力事業団
- 国際協力事業団（2002c）『マラウイ全国スクールマッピング・マイクロプランニング - マイクロプランニング・コンポーネント最終報告書 和文要約』国際協力事業団
- 国際協力事業団（2003a）『エチオピア国オロミア州住民参加型基礎教育改善プロジェクト事前調査報告書』国際協力事業団
- 国際協力事業団（2003b）『日本の教育経験を途上国協
- 力にどう生かすか』国際協力事業団
- 国際協力機構（2004）『教育開発プロジェクト実施・評価体制改善のためのプロジェクト研究報告書』国際協力機構
- 国際協力機構（2005a）『客員専門員報告書 住民参加型小学校建設プロジェクト』国際協力機構
- 国際協力機構（2005b）『世界のよりよい明日のために - JICAのミレニアム開発目標（MDGs）に向けた取り組み - 報告書』国際協力機構
- 国際協力機構（2005c）『マラウイ全国地方教育支援計画策定調査最終報告書 - 和文要約』国際協力機構
- 国際協力機構（2006）『ニジェール住民参加型学校運営改善計画（みんなの学校プロジェクト） みんなの学校便りVol.13』国際協力機構みんなの学校プロジェクト
- 国際協力機構ホームページ（URL: <http://www.jica.go.jp/>）
- 斉藤文彦（2003）『参加型開発』日本評論社
- 佐藤寛（1997）『援助の実施と現地行政』アジア経済研究所
- TICAD市民社会フォーラム（TCSF）（2006）『アフリカ政策市民白書2006 - アフリカ開発と市民社会 - 』晃洋書房
- JICA（2001）. *Final Report of the Regional Educational Development and Improvement Project (REDIP) in Indonesia*, JICA.
- JICA（2002）. *Final Report of the National School Mapping and Micro-Planning Project in the Republic of Malawi: Micro-Planning Component*, JICA.
- JICA（2004）. *Final Report of the Study on the Improvement of Environment for Early Childhood in the Republic of Senegal*, JICA.
- JICA（2005）. *Final Report of the National Implementation Program for District Education Plans (NIPDEP) in Malawi*, JICA.
- JICA（2006）. *Interim Report of the Project on Increasing Access to Quality Basic Education through Developing School Mapping and Strengthening Micro-Planning in Oromia Region, Ethiopia*, JICA.

（2007.2.14受理）

Evaluation of Education Improvement Projects from the Viewpoints of Contribution to Community Empowerment

Yoko Ishida

international Development Center of Japan (IDCJ)

ishida.y@idcj.or.jp

Summary

Various activities with community participation to improve the primary education, such as school construction, have been conducted in the developing countries, which have limited resources for education development. Administrative and technical supports by local governments are essential to make these community efforts effective, including deployment of teachers and delivery of textbooks; however, there are many cases reported that local governments do not provide proper supports. Some of the technical cooperation projects in the education sector, which are implemented through the official development assistance (ODA) of Japan, include (or included) a component to develop local governments and officers' capacity in supporting community participation. These projects are expected to lead to the good governance through community empowerment. Currently, these projects are evaluated from the viewpoints of education development; it might be important to evaluate how they contribute to the enhancement of participatory development at the community and at the government levels.

Keywords

primary education improvement through community participation,
local governments' support to the community participation activities,
promotion of community empowerment and good governance

【研究論文：依頼原稿】

ノンフォーマル教育援助における参加型評価手法の活用 - 「利害関係者が評価過程に評価主体として関わること」の意義 -

源 由理子

国際開発コンサルタント
yminamoto@mvd.biglobe.ne.jp

要 約

近年、貧困削減にノンフォーマル教育の果たす役割が注目される中、ノンフォーマル教育援助ではその活動の多様性や柔軟性から評価が困難であることが課題となっている。本稿は、参加型評価がノンフォーマル教育援助を評価する際に有効に機能するという仮説のもと、その適応可能性を二つの事例を通して考察したものである。参加型評価は、「利害関係者が評価活動に関わる評価」であり評価過程に関わることにより及ぼす影響が多々ある。考察の結果、利害関係者間の「対話」を通して、態度・行動変容といった質的側面の指標化が可能になること、参加型評価の形成的な機能はより適切な提言策定とその活用に関わる可能性が高いこと、関係者が評価の一連の過程に関わることにより評価能力向上を含めたマネジメント能力の向上、ひいては組織強化に関わる可能性があること、が示唆された。他方、参加することへのインセンティブ、中立性の確保、実施体制、評価専門家の資質などに関連する課題もある。本研究の示唆を踏まえ、更なる実践の積み上げを通して手法の検証を行う必要がある。

キーワード

参加型評価、ノンフォーマル教育、質的データ、組織学習

1. はじめに

近年の開発途上国の教育の協力は、1990年にタイのジョムティエンで開催された「万人のための教育会議」において、全世界の子供、青年、成人に対する質の高い基礎教育を提供すること（Education for All: EFA）に対するグローバル・コミットメントを中心に展開されてきた。10年後の2000年にはセネガルのダカールで「世界教育フォーラム」が開催され、2015年までにEFAの達成をめざすべく、「ダカール行動枠組み」が示された¹。

同枠組みでは学校教育機会の拡大がめざされる一方で、EFAの達成のためには学校教育だけではなく、生活技能習得のために青年・成人に対するノンフォーマル教育を充実する必要性も強調された。特にミレニアム開発目標（Millennium Development Goals）の採択により、「貧困削減と成果重視」の戦略が明確に示される中で、貧困削減にノンフォーマル教育の果たす役割に注目が集まっている。

他方、ノンフォーマル教育は「投資の見返りが明確ではない」、「モニタリング・評価体制が確立

されていない」などの理由により、先進国からの援助をひきつけるのが難しいとされてきた分野である（JICA 2005）。ノンフォーマル教育援助の評価に関しては、対象社会・経済の個々のニーズを反映したアプローチの多様性や質的側面の測定困難さなどの特徴から、その難しさが指摘されている。

ノンフォーマル教育に限らず、社会開発セクターでエンパワメントを目指すプロジェクトにおいては、その成果を計る指標の設定が難しく、また効果発現までに時間がかかることからモニタリング・評価の実施とそのための適切な体制構築は常に大きな問題となっている。（中略）特に、活動の多様性と柔軟性、また関係者の多様性が大きな特徴でもあるノンフォーマル教育は、そもそも制度化されていないことが利点であるため、モニタリングや評価は非常に困難である。（JICA 2005：136～137）

ノンフォーマル教育のプロジェクト評価を考えると、近年、開発援助分野において主流となりつつある「結果重視のマネジメント（Results-Based Management: RBM）」によるドナー主導の評価だけでは不十分ではないか、という問いが本研究の出発点である。RBMでは指標化した成果を以って援助予算の資源配分に反映することと、それにより外部アカウンタビリティを確保することをめざしており、ドナー組織のアカウンタビリティを高める上で重要なマネジメント戦略になっている。RBMでは介入による結果を評価の焦点としている。結果（results）は広くアウトプット（output）、アウトカム（outcome）、長期的効果（impact）を包括する概念であるが、個別プロジェクトの評価からプログラム・レベルの評価へとその関心が移りつつある中で、アウトカムやインパクトを重視する傾向へ移りつつある（Binnendijk 2001）。その評価手法は、あらかじめ設定された数値目標に基づきその達成度合いを測定するという「業績測定」の流れを汲み、援助実施の妥当性や短期間の成果として設定された目標の達成度（指標目標値との比較）を測定すること

が中心で、科学的で客観的とされる定量指標を用いている。

ここで気をつけなければならないのは、先進国ドナーによるこのような評価は、援助する側と援助される側、評価する側と評価される側といった非対称性を伴う「援助する側の論理構成に基づいた評価」になりがちであって、被援助側の関心とは必ずしも一致しないことがあるという点である。被援助側の関心は社会経済全体の長期的な開発とそのため介入の極大化であり、期間を限定して実施した援助事業の短期的な成果の評価に対する関心は低い（長尾 2001）。ドナーとして投資効果と外部アカウンタビリティを確保するための評価は不可欠であるが、その一方で、援助評価の結果は開発プロセスの促進や開発効果の持続性に貢献しなければ意味がない。

特に近年増えつつある社会開発セクターの援助では、被援助側の社会・文化的文脈における開発の意味に着目した評価が重要となる。社会開発セクターでは個々の多様化した社会的ニーズに柔軟に対応することが求められる上に、エンパワメントといった人々や社会の変革を支援する要素が強いからである。外からの介入者である援助側は、社会を構成する人々の価値を基準として評価を行う必要があるのではないか。そのためには、援助の非対称性の垣根を取り払い、援助される側の人々を巻き込み評価プロセスを共有することが、事業改善や開発プロセス促進のための有用な評価につながるのではないかと考えられるのである。本稿で対象とするノンフォーマル教育も、「学校教育制度以外の教育」という意味で、個人やコミュニティが置かれた社会の多様なニーズにその都度適応し、識字教育のみならず人々のエンパワメントや社会関係資本の構築・強化までを含むものである。そのようなプロジェクトの評価方法として、実施の過程に評価過程（evaluation process）を組み込み、彼らの開発プロセスを促進する一助となるような評価アプローチ、すなわち「参加型評価」が有効ではないかと考えられる（Minamoto & Nagao 2005）。

プログラム評価の研究では、近年、利害関係者を巻き込んだ参加型評価の理論構築と実践が進みつつある（Patton 1997; Cousins & Whitmore 1998;

Weiss 1998; Fetterman 2001; O Sullivan 2004 他)。参加型評価が登場した背景には、多額な費用をかけた専門家による大規模な評価結果が必ずしも有用な活用に繋がっていないことへの懸念、評価基準が評価対象の固有の状況や背景にある社会・経済・政治的文脈の多様性を反映していないかもしれないことへの疑念がある。参加型評価と一言でいっても、利害関係者評価 (Stakeholder-based Evaluation)、協働型評価 (Collaborative Evaluation)、実用重視評価 (Utilization-focused Evaluation)、エンパワメント評価 (Empowerment Evaluation) など様々な評価理論と方法論があるが、共通する点は評価活動に評価専門家以外の関係者 (= 利害関係者) が「参加」し、評価過程を共有することにより付加価値を高める評価である点だ。「利害関係者」はサービスを受益する側だけではなく評価対象となんらかの利害を有する人々を意味し、たとえば 評価対象プログラムの実施スタッフ、サービスの受益者、資金提供者、政策決定者・意思決定者、評価の資金提供者、当該プログラムと競合する者、当該分野の専門家などがある。どの利害関係者が評価に参加するのが適切なかは評価の目的や使われる方法論によって異なり、プログラムの受益者 (たとえば住民) の参加が前提になるものでは必ずしもない²。従来型評価が評価専門家の出した評価結果による影響を重視しているとしたら、参加型評価は評価過程 (evaluation process) 自体が評価に参加した利害関係者へ与える影響を重視した評価方法であり、評価過程を活用して改善・変革を促す評価であるといえる³。

本稿では、このような参加型評価の、ノンフォーマル教育援助における適応可能性について検討することを目的とする。いかなる評価もその目的によって適応可能な評価手法は異なる。したがって参加型評価の検討は従来型評価の代替として行うものではないということを最初に述べておく。以下、次節では、まず、ノンフォーマル教育援助の特性について整理し、第3節ではノンフォーマル教育に参加型評価を適応した事例を考察し、第4節で結語として今後の実践に向けての示唆を提示する。

2. ノンフォーマル教育援助の特性と評価

(1) ノンフォーマル教育の定義

「ノンフォーマル教育」という言葉が教育政策上に登場したのは1960年代後半から1970年代にかけてとされている (Smith 2005)。その背景には、開発途上国だけでなく先進諸国においても、従来の公的な教育システムが適切に機能していないという危機感があった。1960年代の開発途上国においては、教育は独立以降の近代化のための手段として位置づけられており教育システムの構築は最重要課題の一つであった (黒田・横関 2005)。1960年代後半になると、経済成長の手段としての教育が十分に機能していないという現状を受け、教育が社会・経済の変化に迅速に対応するためには学校教育だけではなくより広い社会的な関わりが必要であるとして、「フォーマル教育 (Formal Education)」、「インフォーマル教育 (Informal Education)」、「ノンフォーマル教育 (Non-formal Education)」の三つの分類が提示された。世界銀行のノンフォーマル教育研究において中心的な役割を果たしたCoombsは、三つの定義を以下のように説明している (Coombs & Ahmed 1974: 8)。

Formal Education :

初等教育から高等教育まで高度に制度化された教育システム。年代順の学年構成や階層的組織を伴うもの。

Non-formal Education :

フォーマル教育システムの外側で行われる組織化された、体系的な教育活動。子どもから成人までの特定のサブ・グループに対する特定の学習内容の提供。たとえば、農業普及・技術訓練、成人識字、フォーマル教育以外の職業訓練、保健・栄養・家族計画・共同組合などの諸分野の様々なコミュニティプログラムを含む。

Informal Education :

家庭、職場、遊びなどの日常の経験や生活の中において知識や技術、態度、考え方などを習得するプロセス。インフォーマル教育は組織的なものではないが、高度に教育された人

たちを含めて、すべての人々の生涯学習に影響を与えている。

Informal Educationは非組織化された一生涯にわたる学習プロセスのことであり、近年、日本語では「インフォーマル学習」と訳されている(たとえばJICA、2005)。したがって、社会における教育は大きく分けて、制度化された正規の教育制度の枠内(=学校教育)で組織的に行われるフォーマル教育と、正規の学校教育制度の枠外で組織的に行われるノンフォーマル教育の二つ分類され、そのすべてのプロセスにおいてインフォーマルな学習が体験されていると考えることができる(Rogers 2004a)。

(2) ノンフォーマル教育援助の分類

「学校教育制度の枠外の組織的な教育活動」であるノンフォーマル教育が開発援助においてカバーする範囲は広く、その形態や種類は多岐にわたっている。たとえば、国際協力機構(Japan International Cooperation Agency: 以下、JICA)によるノンフォーマル教育援助のアプローチを例に取って見てみると、基礎教育の拡充と質の向上(成人・青年の識字率の向上、子どもの識字率・就学率の向上含む)、生計の向上、保健・衛生環境の改善、自然環境保全、平和構築の5つのアプローチに分類され、ノンフォーマル教育を行うことにより期待されるアウトカムが異なるだけでなく、その活動戦略・アプローチは多様であることがわかる。一見、フォーマル教育以外の「教育の要素」を含む開発援助のアプローチをすべて網羅しているようにも見えるが、JICAによると「基礎的な学習ニーズの充足に資する協力を行うこと」がノンフォーマル教育援助の出発点であるという(JICA 2005: 15)。したがって、いずれのアプローチも社会・経済的に排除された弱者(the vulnerable)を対象としていることが特徴的である。

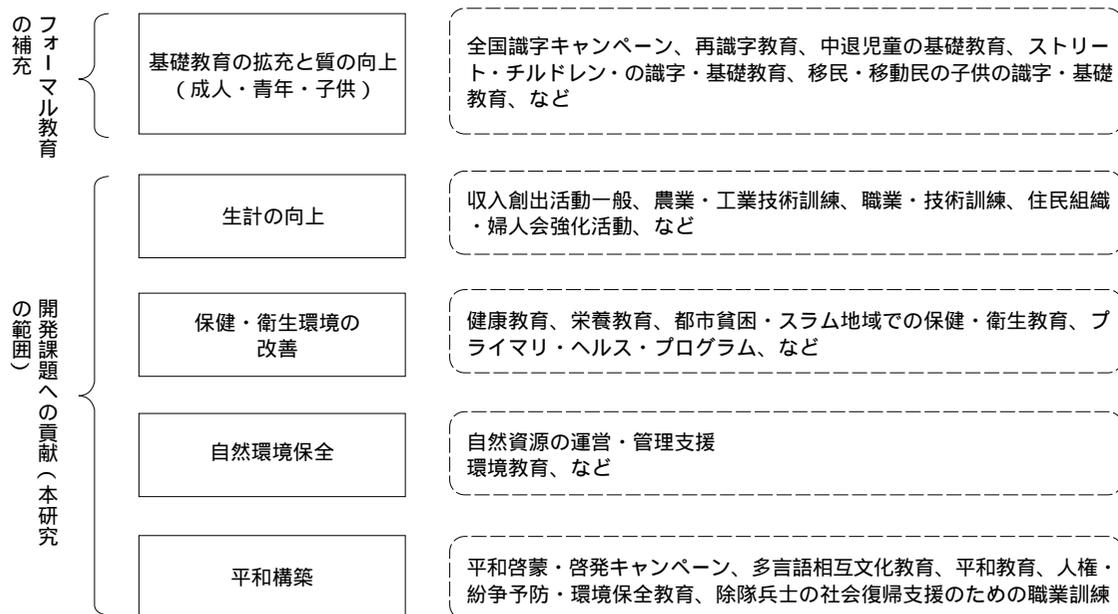
しかしながら、教育の内容に照らし合わせるといくつかの相違があることがわかる。すなわち、「基礎教育の拡充と質の向上」アプローチは不就業児の学校教育への復学や上級学校への進学、成

人の公的職業訓練校などへのアクセスといったフォーマル教育を補完する位置づけが強いが、それ以外のアプローチは、教育を通じて開発途上国が抱える社会的課題の解決や社会変革を図ろうとするものである点が異なる。後者のアプローチでは、対象となる社会・地域の文脈の中で、特定の裨益者のニーズを反映した目標達成の手段として教育が位置づけられているということが出来る。たとえば、貧困層の女性に対するマイクロクレジット・プロジェクトにおける教育、HIV/AIDSの予防教育、住民参加型による公園管理方法の研修、紛争後の帰還・再定住者に対する職業・技能訓練などである。教育と社会との関係で言えば、「基礎教育の拡充と質の向上」が標準化されたカリキュラムにそった既存の社会を維持する人材育成に貢献するものであるとすると、生計向上、環境保全、平和構築などのノンフォーマル教育は、学習を通して個々人の主体性に基づくコミュニティの社会関係資本や価値の形成を促し開発課題の解決をめざすもので、「変革的な志向」が強い。後者はフォーマル教育の補完というよりも、コミュニティや学習者のニーズを反映したもうひとつの教育として切り取ることも可能である(Coombs & Ahmed 1974; Hoppers 2005)。ノンフォーマル教育のモニタリング・評価の難しさは、ノンフォーマル教育の実施体制の柔軟さに起因するものや人々の行動変容にかかる質的側面の評価に関連するものが多いことから、本稿では主に後者のケースを念頭に考察を進めることとしたい⁴。

(3) ノンフォーマル教育の特性と評価の困難性

図1に示すように、教育セクター以外の開発課題の解決を目指すノンフォーマル教育は、ある特定の地域・コミュニティをベースにしたボトム・アップ型の教育である(Smith 2004; Norland 2006)。学校教育と異なり、コミュニティのニーズに基づき、カリキュラムの開発や授業方法について学習者の意向が反映されることが多く、生徒と教師の関係性は上下というよりも水平である。また、教育の実施体制は柔軟である。学習者のニーズや関心を吸い上げるノウハウ(たとえば、仲介者やファシリテーターとしての介在、参加型

図1 ノンフォーマル教育援助の分類と本研究の範囲



(出所) JICA (2005: 17) をもとに筆者作成

アプローチなどによるプログラムの開発)の組み込みと柔軟な仕組みが必要になる。実施期間は比較的短期間で完結するものが多い。教育の内容も標準的なものがあるわけではなく、コミュニティの要望や置かれた状況によって異なり、ライフスキルに直接結びつく実践的なものが多い。たとえば、農業普及技術、生計向上のための手工芸、栄養状態を改善するノウハウなどである。それと同時に、学習者がノンフォーマル教育に参加するプロセスを通して、それら知識・技術のみならず課題解決に向けて主体的に取り組む姿勢を習得することも大きな目的となっており、教師と学習者のインターアクションを生み出す民主的なプロセスそのものが学習の対象となっている (Smith 2005)。またノンフォーマル教育が教育セクター以外のプロジェクトのひとつのコンポーネントになっている場合が多い。たとえば、「貧困支援対策村落開発計画」、「リプロダクティブ・ヘルス・プロジェクト」、「生物多様性保全計画」、「除隊兵士の社会復帰基礎技術訓練プロジェクト」などである。開発援助プロジェクトのモダリティから捉えると、特定のセクターやサブセクター、もしくは

は複数のセクターにまたがる目標を達成するためのひとつの活動として位置付けられているのが特徴的である (Rogers 2004b; JICA 2005)。

表1に示すノンフォーマル教育を評価する場合の困難性を二つの視点から指摘したい。第一に、全体のプロジェクト(あるいはプログラム)もしくはノンフォーマル教育を実施することにより期待される最終アウトカム(あるいはインパクト)にいたる中間的なアウトカム(人々の行動変容などを測定することの困難さがある。たとえば、健康状態が改善されることを最終的に目指すプロジェクトでは、罹患率の低下などのアウトカム指標のほかに、それに至るプロセスで健康教育を受けた母親の健康に関する認識の変化や、食事、出産、日常生活における行動変容を把握する必要があるが、これらの指標をどう取るのがポイントとなる。

第二に、「ボトム・アップ」、「学習者のニーズ」、「柔軟な実施体制」、「主体性の確保」、「問題解決能力の習得」といったキーワードを伴う開発援助事業は、多くの場合、この中間的なアウトカムの指標や目標値を事前に設定する(青写真を描く)

表1 フォーマル教育とノンフォーマル教育との比較

	フォーマル教育	ノンフォーマル教育 ^(注)
アプローチ	・トップ・ダウン ・教師主導	・ボトム・アップ ・コミュニティのニーズ、学習者の意向反映
実施体制	・高度に制度化された体制 ・資源集約型	・柔軟な体制 ・資源節約型
実施期間	・長期間（学校システムへのフル参加）	・一定期間
教師	・有資格者（教員） ・指導者	・認定者、ボランティア ・ファシリテーター
学習者	・標準化されたカリキュラム ・学問志向 ・学習資格がまずありき（学齢）	・個別対応 ・実践的 / 生活密着型 ・学習する人々がまずありき
教師と学習者の関係性	・上下関係	・水平関係
主な目的 / 開発との関連	・学力向上 ・既存の社会システム維持への貢献 ・社会経済開発	・生計向上、生活改善、保健・衛生改善、平和構築、環境保全など ・個人の問題解決能力、主体性の確保 ・社会開発 / 人間開発

(注) ここでいうノンフォーマル教育とはフォーマル教育の補完としてのノンフォーマル教育（例：JICAで分類する「基礎教育の拡充と質の向上」に含まれるもの）を除いたもの。

(出所) Coombs & Ahmed (1974), Smith (2005), 村田 (2005) を参考に筆者作成

というマネジメント手法の限界に直面する (Marsden & Oakley 1990)。しかも、態度変容の指標は「客観的」な定量指標で測定することが難しいとされる (Veenhoven 2002)。しかしながら、ノンフォーマル教育を導入することによる中間的なアウトカムがコミュニティによる持続的な開発を支援する戦略として重要であるとしたら、ブラック・ボックスとするのではなく、その戦略をきちんと評価し事業の改善に生かすこと、すなわち人々、組織、コミュニティの変容と最終アウトカムとの帰属性を評価することが重要であろう。そのためには、コミュニティによる「主観的」な指標を客観的に定量化できる測定方法で収集するなどの検討が必要となる。

他方、参加型評価は、「利害関係者が評価活動に関わる評価」であり、評価の枠組みは関係者自身の興味、関心に反応したものになるため、関係者自身の評価対象プロジェクトに対する認識の変

化や実施のプロセスで起きていることを彼等の文脈の中で解釈することができるとされている。当事者意識の変化など質的な情報が多いことが想定されるが、代替指標 (proxy indicator) や評価基準を関係者で合意するという過程を踏むことにより、「納得性の高い」評価結果を導き出すことが期待できる (Patton 1997)。次節では、このような特性を有する参加型評価の活用事例をとおして、ノンフォーマル教育援助における参加型評価適応の可能性を考察する。

3. 事例

参加型評価を活用したノンフォーマル教育援助の事例として、まず筆者が評価専門家として従事した「ケニア国貧困層エンパワメント・プロジェクト」を取り上げる。二つ目の事例として、アメ

リカ国内のノンフォーマル教育（環境教育）における参加型評価の活用報告を引用する。アメリカの環境教育の事例を取り上げた理由は、第一にノンフォーマル教育援助における参加型評価の事例が極めて限られていること、第二にケニアの事例に限られた利害関係者を参加主体としているのに対し、同事例は利害関係者の多数を評価に巻き込んでいる点で異なった視点による考察が可能になるのではと考えたからである。

(1) 事例1:ケニア国貧困層エンパワメント・プロジェクトの終了時評価における参加型評価⁵

事例概要

当該プロジェクトは、ケニアのNGOであるSave the Children Center（以下、SCC）が従来から進めてきたストリート・チルドレンやスラム住民に対する生活改善事業をさらに推進することを目的とし、JICAの開発パートナーシップ事業として2001年6月から2004年5月まで、また引き続き技術協力事業として2005年6月まで実施されたものである。対象地域はナイロビ市にある4ヶ所のスラム地域 - マザレ（Mathare：人口約30万人）、ミトゥンバ（Mitumba：人口約7,000人）、キベラ（Kibera：人口約50万人）、シティ・カートン（Buru Buru Carton City：人口約2,000人） - である。プロジェクトの目標は、「住民たちの態度や行動変容を通して、自らの力で問題を解決することを通して生活状況が改善すること」であり、そのための主な活動には以下のようなものがある。

主な活動

ストリート・チルドレンや住民に対するエンターポイントの活動。具体的には、サッカー、音楽教室、カウンセリング、ワークショップの実施など

継続的なカウンセリング、リハビリテーションの支援

生活改善ワークショップ、保健衛生ワークショップ、健康ワークショップなどの実施

技能訓練、職業訓練、識字教室などの機会供与
技能、職業、保健、病院等、住民の生活改善に関わる情報の供与

他の関連NGO等の情報提供と住民自らが他の機関と連携していくための支援

このプロジェクトの戦略の大きな特徴は、モノを与える援助は依存症（dependency syndrome）を生み出し自立発展にはつながらないという信念に支えられている点である。したがって、物質的な援助を最小限にし、スポーツ活動、カウンセリング、ワークショップ、基礎的な教育・訓練、機会・情報の提供などを中心として支援を行ってきた。そのプロセスで、住民自らの意思で新しく行動を起こし、問題解決を行う能力を身につけていくことこそ持続的な生活改善への道、と位置づけている。また訓練やワークショップの機会を提供する際にも、住民側からの申し入れや主体性を第一に考え、NGOスタッフがファシリテーターに徹しているということも特徴的である。この背景には住民の貧困は、機会や情報を与えられていないこと、彼らの能力を引き出す環境が与えられていないことに起因するものが大きい、という基本的理解がそこにはある。

このような戦略性を持つ本プロジェクトは前述したノンフォーマル教育の特性を持ち合わせており、特に「生計向上」の事例として取り上げることができる。「生計向上」の定義は、「家族の収入が向上・安定化すること、かつ得た収入が個人の能力あるいは家族の属する地域社会における社会関係資本の構築・強化によって適正に管理・保護されることによって、持続的に生活が改善されている状態」⁶である（JICA 2005：39）。また、JICAによると、生計向上プロジェクトにおいてノンフォーマル教育を取り上げる意義は、1）すべての人々に学びの機会を提供すること、2）様々な生計向上の多様な課題（例：収入向上、情報へのアクセス、ライフ・スキルの習得、機能的識字、会計能力、交渉能力など）に対応できる柔軟性があること、3）人々が自ら課題を認識し、解決方法を見出していく自信やイニシアティブの醸成といった長期的な取り組みを可能にできること、並びに4）持続的な仕組みづくりに貢献する社会関係資本醸成の機会を提供できること、の四つにあるとされている（JICA 2005：39-40）。本事例はこれらの側面を持ち合わせており、評価を

行う際にはとりわけ、「人々が自ら課題を認識し態度・行動変容が起こる」という側面をどのように指標化するのかが大きな課題となった。

評価手法の概要

住民自身による持続的な取り組みが可能になっているかどうかを見るためには、ある一点を捉えて生活状況が改善したことを把握するだけでは不十分である。したがって、生活改善状況を見るといったアウトカム評価だけではなく、プロジェクトの介入を受けた住民の認識の変化や態度・行動変容の変化が起きているかどうかといったプロセスを捉えることが重要であると考えた。その指標を設定するためには、当該住民にとって態度・行動変容とは具体的にどのようなことを意味するのかをまず明らかにしなければならない。それらの意味づけは当然のことながら当該対象社会の文脈によって異なってくる。また、自己決定能力の内容や変化の速度・レベルは受益者（ここではコミュニティ住民、母親、ストリート・チルドレンと多岐にわたる）によって違う。そこで、住民との信頼関係が厚いSCCのプロジェクト・スタッフが「住民の視点」を持つ（実際に対象コミュニティに住んでいるスタッフもいた）と仮定し、彼等の参加を得て指標及び代替指標の検討や評価調査を協働で実施する参加型評価のアプローチを採用した。評価専門家とプロジェクトの利害関係者が一緒になって評価設計から評価結果の解釈までの一連の評価活動を行うものである。本事例の参加者は主にSCCスタッフであり、評価目的は事業の改善である。

まず、NGOスタッフの評価に対する理解を深めてもらうために評価トレーニングを実施し、続いて参加型ワークショップにおけるロジック・モデルを活用したプロジェクトの現状把握と指標検討、並びに評価の設計を参加者全員で行った。態度・行動変容のプロセスはプロジェクト・デザイン・マトリックス（PDM）やロジカル・フレームワーク（ログフレーム）の直線的な因果関係で表すことには必ずしも馴染まないため、想定される因果関係あるいは相関関係を自由に書き込むフロー図を作成した。この方法はプロジェクトの介入を受けて人々が態度・行動変容を遂げていくプ

ロセスとその期待される結果を明らかにし、態度・行動変容の指標を考える上で役に立った。これら一連の作業とともに並行して実施したのは、住民グループに対するフォーカス・グループである。そこでは、住民自身のプロジェクトへの印象、プロジェクトと関わることによる変化、あるいは最近の変化の「きっかけ」などの質的データを収集した。

NGOスタッフとの協働作業と住民のフォーカス・グループにより集められた質的データは、プロジェクトの介入によって起こりうる、あるいは起こったであろうプログラム・セオリーを構築することに活用した。つまり、プロジェクトの質的データを基に帰納的に仮説生成を行うという質的研究の方法を活用したのである。この作業を通して現場で起きている物語（program & story）が明らかになり、数値化された現象の背後にいる人間に光を当て、現象の理解を深めることが期待できる（Patton 2002）。セオリーの構築においては、住民の態度・行動変容の代替指標、最終アウトカムである生活改善の代替指標が特定され、それらを住民に対する質問票調査の選択肢に反映させ、定量化を図った。たとえば態度・行動変容の代替指標としては、「活動に参加するのが楽しくなった」、「友達が増えた」、「自分たちで医療機関にいけるようになった」、「親との関係が良好になった」、「住民グループを自分たちで作った」、「住民グループによる収入活動が始まった」、「コミュニティへの帰属意識が生まれた」、「コミュニティで相互に助け合うようになった」などがある。また、ノンフォーマル教育としてのプロジェクトの活動（スキルズトレーニング、識字教室、生活改善ワークショップ、保健衛生ワークショップ、健康ワークショップなど）が影響を及ぼしているのかも評価する必要があるが、「教育」、「訓練」、「ワークショップ」という言葉を使う代わりに、住民自身が「きっかけ」として感じている事柄をフォーカス・グループの質的データから抽出し、代替指標として設定した。質問票調査結果の解釈は、NGOスタッフとともにワークショップ形式の対話をとおして行った。事業改善を目的としたものであることから、提言の策定も重要である。ワークショップでは評価の分析結果に基づく根拠

に基づき、改善に向けての具体的な取り組みについて議論され、教訓・提言として評価報告書に取りまとめられた。

(2) 事例2：環境保全教育の参加型評価⁷

事例概要

本事例は、アメリカのノンフォーマル教育である環境保全教育の評価を参加型アプローチにより実施した事例報告である。W.I.N. - W.I.N. (Wonders in Nature - Wonders in Neighborhood) と名づけられた学齢児童向け環境保全教育プログラムは1996年からコロラド州デンバー都市圏で実施されているものである。同教育プログラムの目的は、「子どもたちの自然界に対する気づきと理解を深めること」とされている。プログラムの内容は、様々な学習機会を通して野生生物の生態や資源の保全について正しい理解を深めるための、学際的かつ多様なカリキュラムから構成されている。実施主体はデンバー動物園とデンバー市の野生生物所管部局で、グレート・アウトドア・コロラド・トラスト基金が主な資金提供者である。そのほかにも39にのぼるパートナー団体（NPO、政府機関など）が支援を行っており、多くの機関が関わっていることが特徴的である。

評価手法の概要

本事例の報告者は、W.I.N. - W.I.N.プログラムがノンフォーマル教育ならではの特徴を備えているため、学習者を含む利害関係者を巻き込んだ包括的な参加型評価を実施することが有効なプログラムの実施につながると判断した、という（Somers 2005: 43）。すなわち、学習者自身がプログラムに参加した経験を通して自らが主体的に変わり展開していくことに重点を置いており、個々人の生活環境やニーズを反映したプログラムであればあるほど持続的な効果発現につながるという特徴である。ここで使われた参加型評価の手法は、利害関係者の関わり度合いが事例1よりも多い。まず、異なったポジションの利害関係者が関わっている。プログラムの資金提供者、プログラム・マネージャー、プログラム・スタッフ、並びに参加者（学校の校長、教師、子供の家族）からなる

評価委員会（21名）を中心に、評価活動に参加したい関係者に集まりごとに声をかけ、必要に応じて子供たち自身も加わっている。次に、利害関係者は評価過程のすべてのステップに関わっている。すなわち評価実施可能性の検討（evaluability assessment）、ロジック・モデルの構築、評価計画の策定、データ収集方法の検討、データの収集・解釈、結果の報告の一連の評価活動である。参加型評価では、データ収集方法の検討や収集は専門的な知識とスキルが必要となるため評価専門家が担当することが多いが、本事例では委員会を中心に協議が続いた。

もうひとつの特徴は、長い期間をかけているという点である。報告によると、評価計画策定まで4ヶ月、評価終了まで20ヶ月の歳月を要している。これは、評価結果に行きつくまでに、結果以外がもたらす多面的な影響があったということも意味する。たとえば、評価実施可能性評価の時には、「短期的なアウトカム」や「長期的なアウトカム」を検討し、プログラムのミッションと活動の戦略、各活動とのつながりを検討したことにより、関係者のプログラムへの理解を深めることに貢献し、日常業務で追われている関係者がより大きな視点でプログラムを捉えることができるようになったという。それは、この評価活動とは別に、日常業務の活動や新たなプログラム策定を検討する際にも活用されている。あるいは、「評価マインド（ability to think evaluatively）（Patton 2002）」が培われることにより、プログラム・スタッフの評価能力向上（Evaluation Capacity Building: ECB）が図られ、その後のプログラム運営に好影響を与えたという。これらの変化は、評価過程の活用（process use of evaluation）により発現した学習効果として捉えることが可能である。

(3) 事例の考察

これらの事例から、ノンフォーマル教育における参加型評価の適応について、質的データの指標化と、利害関係者の学習効果と組織強化の二つの側面を取り上げ考察する。

質的データの指標化

ノンフォーマル教育援助がコミュニティの多様なニーズに対応し人々や組織、コミュニティの主体性を引き出しているかどうかを評価する上で、参加型評価による質的データが活用できる可能性が高い。ケニアの事例では、プロジェクトを実施するNGOスタッフとともに、住民からのインプットを踏まえ代替指標の設定と評価基準の検討を行っているのが特徴的である。これは、評価基準を社会・文化的文脈の中でより適切な指標を設定しようとした試みであり、住民にとっての「価値」が何であるのかを探ろうとしたものである。従来型評価のように外部の評価者が事前に「科学的に精査された指標」を使って査定（judgment）をするのではなく、対象社会の価値体系の中で構築された意味が「事実（truth）」であるという立場を取り、「対話により合意された指標」を使って評価を行う。これは構成主義を前提とした第4世代評価の理論にも繋がるものである（Guba & Lincoln 1989）。ケニアの事例では、彼らの認識や価値が何であるかを抽出するために参加型ワークショップやフォーカス・グループといった質的データを収集する手段を活用しているが、課題として質的データに基づき選定された指標の正当性をどう担保するのかという点がある。これについてPattonは、「質的データに基づいてセオリーを構成する作業は実験的な場ではなく日常の場で活動している人々のそのままの状態を調査することであり（naturalistic inquiry）、彼ら自身が使う言葉（narrative）を使って彼らにとっての意味を再構成するプロセスを重視しており、当事者である利害関係者間で生み出された理論は、自らに有用で正確な情報入手を可能にする」（Patton 1997）としている。評価過程に利害関係者が主体的に関わるプロセスが、利害関係者にとってより有用な評価を行う上で必要であるというスタンスは、参加型評価の大きな特徴であるといえよう。

ケニアの事例では、評価専門家は終了時評価以外に、その前の2年間にわたり2回、内部評価を行う機会が与えられている。この間の関係者との「対話」をとおして、評価専門家はプロジェクトそのものへの理解を深め、プロジェクト・スタッフは評価に対する理解を深めていった。参加型評

価により利害関係者間で納得がいく指標を設定するまでには、ある程度の時間的な長さによる取り組みが必要になると考えられる。

利害関係者の学習効果と組織強化

W.I.N. - W.I.N.の事例では20ヶ月という長い期間がかかっている。つまり評価結果のもたらす影響よりも、評価過程が関係者に及ぼす影響に重点が置かれていることが推測される。異なった利害関係者が評価活動を通して学習過程を共有するということは、プログラムのセオリーや価値、組織のミッションを対話と熟考を繰り返して問いかけることであり、評価対象のプログラムや組織、並びに個人の人々の立場に対する理解を深めることにつながる（Preskill Suckerman & Amthrews 2003）。参加型評価の構成主義の論理をここでも読み取ることができるが、そのプロセスで利害関係者が対話を通して学習し合うことが大きな利点として挙げられ、ノンフォーマル教育のひとつの手段としても機能するのではというSomersの指摘は興味深い（Somers 2005:43）。

またW.I.N. - W.I.N.の事例で指摘されているように、長期間の関わりは、利害関係者の評価能力を向上させ、プロジェクトの運営にプラスの影響を及ぼす。それは実施組織・体制の強化にもつながった。ノンフォーマル教育はNGOや住民組織など脆弱な組織が実施することが多いためにデータ収集等を含めたモニタリング体制に大きな問題があることが指摘されているが（FAO & UNESCO 2005）、それを補完する働きが期待される。

さらに、参加型評価が形成評価に適しているという特徴に注目したい。フォーマル教育の高度に制度化された体制と異なり、ノンフォーマル教育は多様な学習者を対象に柔軟に対応していかなければならない。ケニアの事例では、評価に参加したプロジェクト・スタッフから今後のプロジェクト運営（住民へのワークショップ、トレーニングの内容を含む）に関し改善のための具体的な提案がなされた。たとえば、プロジェクトに関わった住民の主体性の表れとして活発化したグループ活動が挙げられるが、さらに自立に向けての機会を提供するためには、コミュニティの外にある情報やサービスに自分たちからアクセスしていけるよ

うな交渉術のトレーニングが必要である、という提案があった。ノンフォーマル教育では、このような学習者の「変化」や、同じコミュニティ内においても異なるニーズに対し対応していくという弾力性を伴うだけに (Coombs & Ahmed 1974 ; Norland & Somers, ed., 2005) 形成的な評価を導入することは事業改善と組織マネジメントの向上に効果的である。

4. 結語：実践への示唆

本研究では、ノンフォーマル教育援助における参加型評価適用の可能性について、ふたつの事例を基に考察した。本稿で念頭においたノンフォーマル教育は、農業開発、農村開発、保健・医療、環境保全、平和教育といった開発課題に貢献するひとつの手段として位置づけられ、したがって考察結果はノンフォーマル教育のみならず、広くそれら開発課題に関する事業評価にも応用できると考えられる。また、参加型評価の適用可能性を検討することは、「利害関係者が評価過程に評価主体として関わること」によって及ぼす影響を検証することでもある。

結語として、まず期待される利点を三点挙げたい。第一に、参加型評価を導入することにより、ノンフォーマル教育の中間的なアウトカムである人々、組織、コミュニティの態度・行動変容といった質的側面を指標化することができると考えられる。質的データを関係者とともに収集し、関係者間の「納得」のもとに設定された指標は、対象社会の価値体系の中で構築された「妥当」な指標として捉えることが可能である。

第二に、ノンフォーマル教育ではコミュニティのニーズや学習者の変化に合わせて教育の機会を提供していくことが重要であるが、参加型評価は形成的な評価に適しており、当事者としての利害関係者自身による具体的かつ実施可能性の高い提言が策定されることが期待できる。また、意思決定を行う人々が評価過程に関わっているため、評価の提言が活用される確率が高い。

第三に、評価過程を共有することによる学習効果の側面に注目したい。参加型評価では評価に参

加する利害関係者へのトレーニングを初めとして、評価可能性評価、ロジック・モデルの構築、評価設計、評価の解釈など一連の評価活動を一緒に行うことになる。その評価過程で 評価対象への理解促進、 評価能力向上、 プロジェクトの運営能力向上が期待できる。ノンフォーマル教育は、フォーマル教育と比較して実施体制が脆弱であるが、参加型評価をとおして実施組織の強化にも繋がると考えられる。また、学習者も評価過程に参加することができれば、問題解決能力や自己決定能力の向上へ繋がる可能性も大きく、教育の手段として位置づけることができる。

他方、利害関係者が評価過程に参加することは、従来型評価と異なる実施体制が必要となる。また、自己評価に近い形で行われるため主観性やバイアスにどう対処するのかという課題にも直面する。最後に、参加型評価を実践する上での課題と留意点、並びに提案として五つの点を提示したい。

まず、参加型評価を行う前提として、プロジェクトの現場にある程度の権限が委譲されている必要がある。現場で事業改善のための意思決定が行われることが多くの参加者を継続的に巻き込むインセンティブにつながることを考えると、ボトム・アップの意思決定の仕組みが必要である。参加型評価は分権化が基本なのである。

第二に、参加型評価はコミュニティや人々の生活に密接に関わりながら作業を行うため、データの収集や評価の判断にバイアスがかかりやすいという批判を真摯に受け止める必要がある。それらバイアスを取り除くためには、評価専門家の果たす役割は大きい。評価専門家は、参加者の一員となることによって「当事者性」をある程度保ちながらも「部外者としての中立性」の視点を持ち、評価活動を適切にファシリテートすることが求められる。質的評価に従事する評価専門家は人々との距離を縮めない限り、妥当な質的データの収集や関係者間の建設的な対話は難しい。Pattonが言う「共感的中立性 (empathetic neutrality)」(Patton 2002 : 50) を保ちつつ、判断を曇らせる感情移入を抑え、理解を鈍らせる距離の置き方をしない、という評価倫理を守る必要がある。

第三に評価に費やす時間の問題がある。参加型評価は、継続的に、ある程度の時間をかけて行う

ことに意味がある。プロジェクト・サイクルの一定の時期に、その都度結成された評価ミッションが評価を行うのではなく、参加型評価そのものを事業の中に内在化する仕組みが効果的であろう。たとえばODAの場合は、参加型評価を行う同一の評価専門家をプロジェクト・サイクルにわたって断続的に派遣するという仕組みが考えられる⁸。またNGOの場合は、評価予算そのものが限られているという問題があるため、助成金の中に参加型評価のコンポーネントを入れることが考えられる。

第四に、異なった利害関係者間の対話を基本とする参加型評価の実施プロセスと結果は、評価専門家の資質に大きく左右されるという点に留意が必要である。評価専門家には、従来型評価に必要な評価技術に加え、評価訓練を行うトレーナー、ファシリテーター、内部のアドバイザーなどの機能を担えるスキルが必要である。特にノンフォーマル教育の評価では、事業の特性として事前にプログラム・セオリーを想定することが困難なケースが多いため、評価の段階において評価可能性評価、ロジック・モデル作成を利害関係者と一緒に行うプロセスが重要であり、評価理論のトレーナー・教師としての役割が大きい(Monroe, M. C. 他 2006)。

最後に、参加型評価は一義的には形成評価による事業改善を目的としたものであり、ノンフォーマル教育戦略の策定やドナーのアカウントビリティ確保が目的の評価、あるいはプロジェクトの有効性を問う評価とは異なることを踏まえる必要がある。インパクト評価を行う場合には、参加型評価を通して収集した指標データを共有することが可能であろう。

評価手法の研究は、その手法を使って実施した評価が他の手法よりも有用か、妥当かの検証も必要である。本研究の示唆を踏まえ、今後も更に参加型評価の実践を積み上げることにより手法の検証を行っていきたい。

謝辞

投稿原稿に対し有益な指摘をくださった匿名の査読者に、記して感謝の意を表します。

注記

- 1 「ダカール行動枠組み」の目標は、就学前保育・教育の拡大と改善、2015年までに無償で質の高い教育をすべての子どもたちに保障すること、青年・成人の学習ニーズの充足、2015年までに成人識字率(とくに女性)を50%改善すること、2005年までに初等・中等教育における男女の格差を解消すること、読み書き、計算および基本的な生活技能習得のために教育の質を改善すること、である。
- 2 本稿では評価対象として開発援助分野の事例を扱っているが、参加型評価の理論は開発援助で使われる「参加型開発」の言説とは異なるものである。
- 3 参加型評価の分類には、その目的、ならびに評価専門家と評価に参加する利害関係者の間の関係者性に着目し、「実用型の参加型評価(practical participatory evaluation)」と「社会正義や弱者の自立を目指す変革型の参加型評価(transformative participatory evaluation)」という分け方もある(Cousins & Whitmore 1998)。
- 4 Rogers (2004b) は、ノンフォーマル教育は教育省による教育そのものを目的にしたものと、教育省以外の省庁による開発課題の解決を目的としたもののふたつが混在しているとした。
- 5 事例の記述は、Minamoto (2005) 及び現地でプロジェクト・スタッフとともに実施したワークショップの記録による。
- 6 英国開発援助庁(DFID)による「生計の自立」(Sustainable Livelihood)に基づく概念である。
- 7 Somers (2006) の評価報告に基づく。
- 8 新たな仕組みを作ることによるコストの問題もあるが、Somers (2006) によると参加型評価は、業務の中に内在化されることにより、大規模なインパクト評価を行う際に雇用する評価専門家に対する報酬を低く抑えることにつながり効率的だという。

参考文献

- 江原裕美編(2003)『内発的發展と教育～人間主体の社会変革とNGOの地平』、新評社
- 黒田一雄・横関祐見子(2005)『国際教育協力の潮流』、『国際教育開発論～理論と実践』(黒田一雄・横関祐見子編)、有斐閣

- 国際協力機構国際協力総合研修所〔JICA〕(2005)『ノンフォーマル教育支援の拡充に向けて』
- 長尾真文(2001)「援助における評価の目的と活用方法：南アフリカ理科教育支援事業による例示」、『国際教育論集』4(1)：89-100、広島大学教育開発国際協力研究センター
- フリック、ウヴェ(2002)(小田博士他訳)『質的研究入門 人間の科学のための方法論』春秋社(原題：Qualitative Forschung by Uwe Flick, 1995)
- 源由理子(2003)「エンパワメント評価の特徴と適用の可能性～Fettermanによる「エンパワメント評価」の理論を中心に～」、『日本評価研究』3(2)：70-85
- 村田敏雄(2005)「識字・ノンフォーマル教育」、『国際教育開発論～理論と実践』(黒田一雄・横関祐見子編) 有斐閣
- Binnedijk, A. (2001). *Results Based Management in the Development Co-operation Agencies: A review of experience*. Background Report. Organization for Economic Cooperation and Development, Paris.
- Coombs, P. H. & Ahmed, M. (1974). *Attacking Rural Poverty ~ How Nonformal Education Can Help ~*. A Research Report for the World Bank. The Johns Hopkins University Press.
- Cousin, J. B. and Whitmore, E. (1998). Framing Participatory Evaluation. *Understanding and Practicing Participatory Evaluation, New Direction for Evaluation*. American Evaluation Association, San Francisco: Jossey Bass, 5-23.
- FAO and UNESCO (2005). *Education for Rural People to achieve EFA and the MDGs*. Background Paper for the Panel Session on the Education for Rural People (ERP) Flagship Partnership. Paris.
- Fetterman, D. M. (2001). *Foundation of Empowerment Evaluation*. Sage Publications.
- Guba, E.G. and Lincoln, Y. S. (1989). *Fourth Generation Evaluation*. Sage Publications.
- Hoppers, W. (2005). *Summary Report on the Electronic Discussion Forum on "Challenging Non-formal Education in Africa: Where do we go from Here; What Role for an ADEA Working Group?"*. Association for Development Education in Africa.
- Marsden, D. and Oakley, P. (1990). *Evaluating Social Development Projects*. Development Guidelines No. 5, Oxfam UK and Ireland.
- Minamoto, Y. (2005). *Completion Report on Expert Assignment Program for Evaluation Capacity Building, EPC*, JICA Nairobi Office.
- Minamoto, Y. and Nagao, M. (2006). Process Considerations in Evaluating Educational Cooperation Projects. *Journal of International Cooperation in Education*, 9(1). 89-105.
- Monroe, M. C., Fleming, M. L., Bowman, R. A., Zimmer, J. r., Marcinkowski, T., Washburn, J. & Mitchell N. J. (2006). Evaluators as Educators: Articulating Program Theory and Building Evaluation Capacity. *Evaluating Nonformal Education Programs and Settings*. New Direction for Evaluation. American Evaluation Association, San Francisco: Jossey Bass, 5-12.
- Norland, E. (2006). The Nuances of Being "Non": Evaluating Nonformal Education Programs and Settings. *Evaluating Nonformal Education Programs and Settings*. New Direction for Evaluation. American Evaluation Association, San Francisco: Jossey Bass, 5-12
- O'Sullivan, R. D. (2004). *Practicing Evaluation ~ A Collaborative Approach ~*. Sage Publications.
- Patton M.Q. (1997). *Utilization-Focused Evaluation*. The New Century Text, 3rd edition. Sage Publications.
- (2002). *Qualitative Research & Evaluation Methods, 3rd edition*. Sage Publications.
- Preskill, H., Suckerman, B. & Matthews, B. (2003). An Exploratory Study of Process Use: Findings and Implications for Future Research. *The American Journal of Evaluation*, 24(4), 423-441.
- Rogers, A. (2004a). Looking again at Non-formal and Informal Education-Towards a New Paradigm. *The Encyclopedia of Informal Education*. http://www.infed.org/biblio/non_formal_paradigm.htm (2006年10月20日アクセス)
- (2004b). *Non-Formal Education ~ Flexible Schooling or Participatory Education?* The University of Hong Kong.
- Smith, M. K. (2005). Non-formal Education. *The Encyclopedia of Informal Education*. <http://www.infed.org/biblio/b-nonfor.htm> (2006年10月20日アクセス)
- Sommers, C (2006). Evaluation of the Wonders in Nature-Wonders in Neighborhoods Conservation

Education Program: Stakeholders Gone Wild!, *Evaluating Nonformal Education Programs and Settings*. New Direction for Evaluation. American Evaluation Association, San Francisco: Jossey Bass, 29-55.

Weiss, C. H. (1998). *Evaluation, 2nd ed.* Prentice-Hall, 1998.

(2007.2.14受理)

Utilizing Participatory Evaluation in Non-formal Educational Cooperation Projects – Relevance of Stakeholders Involvement in Evaluation Process –

Yuriko Minamoto

International Development Consultant
yminamoto@mvd.biglobe.ne.jp

Abstract

Recently, the impact of non-formal education on poverty reduction has been increasingly recognized among international aid communities. However, the difficulty of evaluating non-formal education due to the elastic dimensions of implementation process has become challenging issue. This article discusses applicability of ‘ Participatory Evaluation ’ in non-formal settings through case analysis, assuming it would effectively function in the evaluation process. Various influences on projects could be observed through stakeholder involvement in Participatory Evaluation practice. The following points are concluded as expected effects of its utilization: 1) identification of quality indicators of attitude and behavioral changes through ‘ dialogue ’ among stakeholders, 2) appropriate and practical recommendation that lead to more utilization, due to the formative aspect of Participatory Evaluation, and 3) the contribution to strengthening evaluation capacity as well as organizational learning. More practice and cases of Participatory Evaluation should be accumulated and analyzed for further examination and improvement of its methodology.

Keywords

participatory evaluation, non-formal education,
qualitative data, organizational learning

【研究論文：依頼原稿】

初等教育分野におけるプログラム・ベースド・アプローチの援助に関する評価分析 ホンジュラス及びベトナムのEFA-FTIの事例から

三輪 徳子

国際協力機構

miwa.satoko2@jica.go.jp

要 約

近年、初等教育分野ではプログラム・ベースド・アプローチ（PBA）による援助が活発化している。本研究では、ホンジュラスとベトナムの「万人のための教育 ファスト・トラック・イニシアチブ」（EFA-FTI）を事例に、開発成果実現の観点からPBA援助の形成評価を行った。その結果、アプローチは同じでも開発成果実現に向けた進展には差異があり、計画の適切性、援助依存度、実施能力、実施環境等が影響することを明らかにした。PBAは一律に議論されがちであるが、各国固有の状況をふまえた進め方が必要である。そのためには、対象国の課題をシステミックに捉え、実施能力の強化を盛り込んで段階的に取り組んでいくことが重要であり、イニシアチブやアプローチがドナー主導とならないことが必要である。また、PBAにおける個別援助の貢献は、投入やアウトプットとの関連に焦点を当てて検討するだけでなく、全体の中での位置付け・役割に照らし、実施基盤や実施プロセスへの貢献を含めて評価することが必要である。

キーワード

プログラム、開発成果、援助効果、貢献、キャパシティ・デベロップメント

1. はじめに

(1) 問題提起

プログラムは、「特定の目標を達成するために整理された一連のインターベンション」（DAC 2003）と定義される。開発援助分野のプログラムには、途上国による開発プログラムとドナーによる援助プログラム（国別、課題別の協力プログラム等）がある。現在、我が国ODAで進められているプログラム・アプローチは、特定ドナーが協

力目標の達成に向けて事業をより戦略的に組み合わせさせてプログラム化を図るものであり、プログラムは援助プログラムを指す。これに対してプログラム・ベースド・アプローチ（PBA）は、開発目標の達成に向けて途上国の開発プログラムをドナーが協調して支援するものである。

PBAは、1990年代後半以降の国際社会における、開発成果達成を重視する議論・取り組みをアプローチとして統合したものであり、途上国のリーダーシップの重視、途上国の単一・包括的プログラムと予算枠組みに基づく援助、援助協調・

調和化、途上国制度の活用を要素とする。PBAの一つがセクターワイド・アプローチ (SWAp) である。(CIDA 2003、DAC 2006)

特に教育分野では、初等教育普遍化等が国際開発目標に掲げられた優先分野であることから、他の分野に先行する形で取り組みが進んでおり、PBAが主流化してきている。本論のPBAへの関心は、こうしたPBAの重要性にある。

では、PBAは開発成果の実現にどの程度有効なのか。同視点から、開発成果に向けたPBAのロジックを整理したのが図1である。4つの要素への取り組みを通じ、途上国側コミットメント及びドナーの体系的支援、途上国側開発マネジメントの改善を確保することで、プログラムの適切な実施と開発成果の実現を目指すものとまとめられる。

PBAについては、事例研究を含めて少なからずの先行研究がある¹。これらの研究では、PBAは各国同じようなメニュー（中期支出計画策定、財政支援、ドナー会合制度化、合同評価等）で進められていること、アラインメントや援助協調の進展に寄与しており、特に援助依存度の高い国では有効であること、援助協調等への対応から、調整コストは短期的には必ずしも低減していないこと、プログラムの進捗は国によって異なり、対象国の経済運営、コミットメント、実施能力等が影響すること、などが示されている。

しかし先行研究は、PBAが開発成果の実現につ

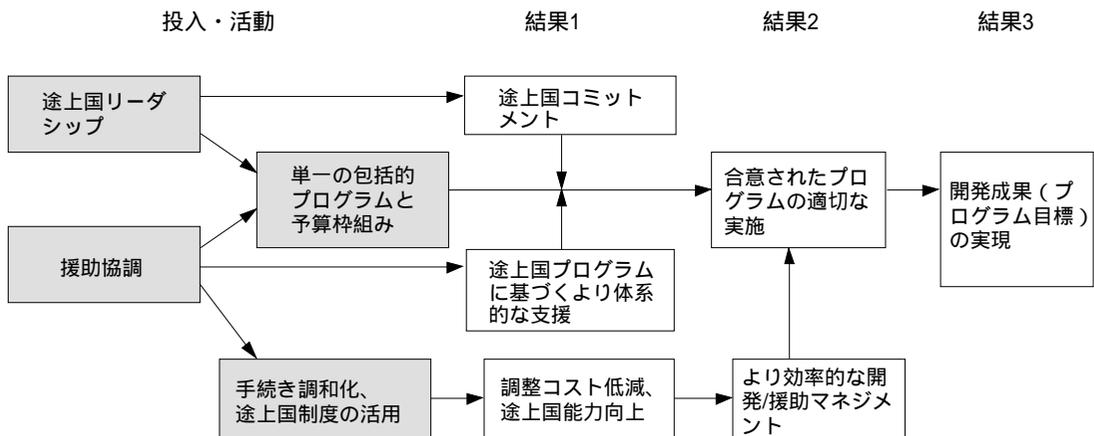
ながっているのかには言及していない。事例の多くが進行中との事情はあるが、PBAの進捗による開発成果の実現可能性もあまり触れられていない。大半の研究は、図1の「投入・活動」あるいは「結果1」までの進捗及びその影響要因、すなわちプログラムへの取り組み方の分析に留まっている。開発成果の実現可能性の観点から具体的にプログラムの内容やドナーによる支援内容を分析したようなものは見当たらない。

(2) 本研究の目的、方法

本論の目的は、先行研究の上記状況をふまえ、PBAが進んでいる「万人のための教育 ファスト・トラック・イニシアチブ」(EFA-FTI)を取り上げ、「結果2」及び「結果3」に対する評価を試みることで、PBAの今後のあり方について示唆を得ることにある。EFA-FTIは、「万人のための教育」への取り組みの促進を目的に2002年に導入された国際イニシアチブである。対象国を選び、同国のEFA-FTIプログラムの実施をPBAにより国際的に支援している。

評価では、EFA-FTIの先発対象国であるホンジュラス及びベトナム（2002年及び2003年に対象国に承認）を事例として、評価調査時点（ホンジュラス2005年、ベトナム2006年）までの進捗をもとに、以下の評価設問により分析した。

図1 PBAの概念



(出所) Bartholomew et al 2006、CIDA 2003、OECD/DAC 2005、Walford 2003をもとに筆者作成

[評価設問]

両国EFA-FTIは開発成果実現に向けて、適切な実施が進んでいるか。

プログラムは、開発成果実現に必要な計画の適切性を満たしているか

途上国のコミットメントとドナーの援助協力が、開発成果実現に必要な体系的なプログラム介入に繋がっているか

プログラムの実施プロセスは円滑か

両事例から導かれる、PBAに関する今後の留意点は何か。

なお本研究は、両国の初等教育分野におけるJICAプログラムを対象とした評価調査での収集データに基づいている（ホンジュラスは2005年5月、ベトナムは2006年6月に現地調査実施。筆者は両調査に团长として参加）。本論は、上記評価調査では特定ドナー（JICA）のプログラムに限定していた範囲を、対象国の開発プログラム並びに対象国及びドナーによる取り組み全体にスコープを広げ、EFA-FTIの目標達成に向けた進展を評価するものである。

以下、第2節及び第3節ではホンジュラス及びベトナムの各事例に関し、上記評価設問 について検証する。そして第4節で両事例を比較しつつ、評価設問 について考察する。

2. 事例1：ホンジュラスEFA-FTI

ホンジュラスのEFA-FTIは、援助依存度が高い途上国の事例として、援助依存度と途上国主体のプログラムの策定や、プログラムへの体系的介入、途上国の実施能力と、途上国のリーダーシップや途上国制度活用との関係、政治・社会面を含む対象国の実施環境の問題などについて、PBAのもとでの開発成果に向けた取り組みに関する多くの示唆を示している。

(1) 計画の適切性

ホンジュラスでは1990年代以降、教育セクター開発が重視されてきた。この結果、初等教育²の

粗就学率は97%に達しているが、終了率は69%と低い。男女格差は少ないが、所得間・地域間格差は大きく、学齢児童中の非就学人口は都市部4%に対して農村部では10%、平均在学年数も都市部6.7年に対し、農村部は3年の状況にある。先住民族では、学齢児童期の36%が未就学とされている。

これらの要因としては、就学前教育就学率の低さ（39%）、留年率の高さ（6年間での終了率32%）、退学率の高さ（3%）、授業時間の少なさ（年間200日の規定に対し実際は110日）、教員の質の低さ、複式学級の多さ（44%の学校）、農村部における学校へのアクセス、地域間経済格差や貧困家庭における教育への関心の低さなどが挙げられる（Honduran Ministry of Education 2002）。

以上のような課題に対し、EFA-FTIは 初等教育終了率100%、12歳（留年なし）での終了率85%、算数・国語（国際比較で中米諸国中最下位にあり留年等の要因）の学力の向上を全体目標とし、5つのコンポーネントで構成されている。各コンポーネントの目標・内容は表1のとおりである。

プログラムの特徴を見てみたい。まず、就学率は、就学前教育は目標に含まれているが初等教育は明示的には掲げられていない。内容も学校内の教育サービス改善が中心であり、経済格差や教育への関心の低さなど、学校外要因への取り組みはほとんど盛り込まれていない。また、サービス改善には、質とともにアクセス（農村部や少数民族）への取り組みも含まれているが、施設の量的拡充は入っていない。以上からは、EFA-FTIは初等教育普遍化のための課題解決に向けた包括的なプログラムというよりは、学校レベルの教育サービスの質の向上にフォーカスが絞られたプログラムと言える。

ホンジュラスでは、EFA-FTI策定時点での初等教育に関わる開発戦略として、PRSP（2001年）並びに、政府開発計画及びそのもとでの教育戦略・アクション計画（いずれも2002年）があった。EFA-FTIはこれらをベースとしているが、他の戦略では就学率が主たる目標となっており、EFA-FTIには入っていない学校施設改善や教育行政能力の強化が盛り込まれている。EFA-FTIの限定性

表1 ホンジュラスEFA-FTIプログラムの概要と政府・ドナーの取り組み

コポーネント	2015年までの指標	主な内容	ドナー	支援内容
1. 基礎教育の効率性 (6歳で1年生入学 6年間で6年生終了)	・12歳児6年生終了率85% ・6年生終了率100% ・6年生学カテラスト得点率 (算・西) 70%	・1年生への入学 学習基準や支援教材	世界銀行 政府 JICA 米国 カナダ スウェーデン UNICEF	農村部・貧困地域への代替教育 (Comunitaria) 新カリキュラムの策定 ラジオを用いた参加型算数学習 算数の教材作成 学習標準・共通テスト開発 (MIDEH) ラジオによる代替教育 (EDUCA TODOS) ラジオ教育による算数能力向上 (APREMAT) 算数の教材印刷 スペイン語・算数の教材印刷 低学年の読み書き能力強化
2. 質と効率性を伴った 教員 (教員研修と教員 パフォーマンスの 質と効率性の改善)	・授業実施日数200日 (年間1000時間) ・大学卒業教員3000名 ・教員支援センター1500箇所 ・学校・教育の管理システムの強化	初率的な新旧 適齢以上の児童への対応 内部効率性のM&E 中退児童の救済 新規教員養成 教員配置 現職教員研修 教員へのインセンティブ 学校と教員のマネジメント 視学・フォローアップ	政府 ドイツ 世界銀行 JICA スペイン ドイツ 米国 世界銀行 UNICEF スウェーデン 政府 JICA	未修了者対象のラジオ教育 新規教員養成への資金援助 (PRODES) 教員養成校の施設整備 大学への機材供与、計画策定支援 (Comunitaria) 教授法の研修 JICA 算数の現職教員研修 (PROMETAM) 算数の教材活用案に閉する協力 (Luis Lande) 教材開発 教員研修を通じた西語・理科支援 (FEBLI) 新規教員養成研修 (PRODES)
3. 就学前教育の強化 (5歳児の就学前教育の 普及化)	・5歳児就学率100%	カバレッジ 教材 教員・ボランティアへの研修 視学・フォローアップ データベース構築 異文化二言語教育の制度化 学習環境の適正化 教員の研修とパフォーマンス コミュニティ参加 特殊教育	米国 世界銀行 UNICEF スウェーデン 政府 JICA	視学官研修 (Salvemos) 学習標準・共通テスト開発 (MIDEH) 農村部・貧困地域への代替教育 (Comunitaria) ノンフォーマル就学前教育の組織化・休職 読み書き能力改善と地域参加支援 (Escuela Amiga) ラジオ学習の教材開発 教材の調達 機材の調達 ボランティア指導者の研修 幼稚園教員への研修 (モデルプロジェクト)
4. 多文化二言語教育の 公正・アクセシ (二言語異文化教育の アクセシと公正の保証)	・5歳児就学前教育就学率100% ・12歳児6年生終了率85% ・6年生終了率100% ・3・6年生学カテラスト得点率 (算・西) 70%	視学・フォローアップ データベース構築 異文化二言語教育の制度化 学習環境の適正化 教員の研修とパフォーマンス コミュニティ参加 特殊教育	政府 世界銀行 UNICEF	異別技術支援チームの研修、児童へのインセンティブ 特殊教育学校への機材供与 (Comunitaria) 二言語異文化教育支援
5. 農村部の教育ネット ワーク (農村部の5・15歳への教 育アクセス確保のため の教育ネットワークの 構築)	・5歳児就学前教育就学率100% ・12歳児6年生終了率85% ・3・6年生学カテラスト得点率 (算・西) 70% ・ネットワーク学校466校	ネットワーク設立 ネットワークの教授法モデル 連級に対するサポート ネットワークの管理と監督 学校給食・バウチャー ネットワーク評価と情報システム 上記以外 教育者の能力強化支援	政府 スペイン 世界銀行 ドイツ WFP	中核教員グループの組織化 特殊教育分野での技術・資金協力 ネットワーク校への教材配布 (Comunitaria) 技術支援 (2県) 学校給食 貧困地域児童への学習用品配布 (コモニアンド) EFA支援・教育改善支援 教育者の組織強化支援・機材供与 教育者への機材供与 EFA関連部署への技術支援 EFA関連部署への機材供与・スタッフ給与支払い コモニアンド (署名済み) コモニアンド (署名済み) コモニアンド (署名済み) コモニアンド (署名済み) コモニアンド (署名済み)
	財政支援 (コモニアンド)	財政支援 (コモニアンド)	スウェーデン ドイツ カナダ 米国 世界銀行 カナダ 世界銀行 スペイン	

(出所) JICA, 2006をもとに、筆者作成

は、他の戦略との比較でも明らかである。

こうしたEFA-FTIの限定性のもと、特にEFA-FTI実施に伴い教育行政能力の問題が顕在化し、同強化を盛り込んだより包括的な戦略が必要となったことを受け、セクター・プログラムの策定も検討されている³。EFA-FTIや同セクター・プログラムは国際合意に基づく計画である。他方、上記政府開発計画や教育省戦略・アクションプランは国際合意によらない国内計画である。ホンジュラスでは、4年ごとの政権交代に伴い政府開発計画等が新たに策定しなおされることもあり、国内計画よりも国際合意に基づく計画が開発のベースとなっている。セクター・プログラムの検討は、EFA-FTIの限定性ととも、ホンジュラスのこうした状況も背景にある。

EFA-FTIは、ドナー等利害関係者の合意を得て策定されることになっている。にもかかわらず、計画の適切性の点で問題なしとし得ないような内容になったのはなぜか。その一つとして挙げられるのが、EFA-FTIの策定プロセスの問題である。ホンジュラスのEFA-FTIプログラムは、2002年6月の対象候補国発表の後、10月にプロポーザルを事務局提出、11月に現地ドナーによる集中アセスメントと修正作業が行われ、同月には対象国として承認されるといったように、非常に短期間で策定された経緯がある（JICA 2006）。ホンジュラスでは、教育分野のドナー会合（Mesa Redonda de Cooperantes en Educacion（MERECE））が活発に活動しているが、上記のような経緯から、プログラム策定プロセスへのMERECE参画は短期間の上記集中アセスメントに限られている。よって利害関係者の合意は形式的とも言え、PRSPで行われたような市民社会等の参画が実施されていないとの問題指摘もある。こうしたプロセスは計画内容と無関係とは言えない。

もう一つが実効性であり、対象国承認に向けて、実現可能性が高く比較的结果の出やすい活動にフォーカスを絞って策定された面も見受けられる。このうち実現可能性に関しては、援助依存度の高いホンジュラスでは援助の有無が重要な鍵となる。以下（2）に詳述のとおり、EFA-FTIのフォーカス分野は以前から多くの援助が行われてきた分野と合致する。計画策定にあたり課題のみでな

く、こうした面が考慮された。

（2）介入の進展

EFA-FTIの実施に必要な資金源は政府予算とドナーの援助である。このうち前者は2005年で、予算中最大の31%が教育分野に、そのうち30%が初等教育に充てられており、優先分野となっている⁴。教育分野予算の出所も政府予算85%、外国融資3%、無償資金供与8%、HIPC債務削減4%となっており、政府予算が大きな比重を占めている。しかし政府予算の内訳を見ると、97%は経常支出（政府予算平均は69%）であり、その大半は一人当たりGDPの6倍と国際的にも高い教員給与などの人件費である⁵。このため開発事業費は援助に依っており、同年のEFA予算も85%が無償資金供与、11%が外国融資で、政府予算は4%に過ぎない。

こうしたことからEFA-FTIは、財政支援を中心とする援助の増加（資金計画では3年毎の各サイクル、約86百万ドル）を見込んで計画が立てられている。援助モダリティも、2003年10月署名の政府・ドナーによる枠組み合意（General Framework Memorandum of Understanding）において、プール・ファンド、プロジェクト、人や物によるイン・カインド支援、補完プログラム（EFA-FTIには含まれていないが補完性を有する活動）のいずれも可としている。

財政支援は、2004年6月にコモン・ファンドに関する合意（Memorandum of Understanding concerning Pooled Fund Support）が署名され（世銀ほか3ドナー）、ファンドが設置されたものの、教育省の信用管理能力の問題から拠出が遅れており、2005年末までの拠出はスウェーデンの450万ドルのみとなっている。

したがって、援助はプロジェクトが中心である。これらのEFA-FTIへの介入状況（2002年から2005年5月時点までの初等教育分野援助）は表1のとおりである。EFA-FTI開始から2年余りということもあるが、多くの援助が入り様々な取り組みが行われている分野と、そうでないものが分かれている。前者としては質の向上に関する学習基準・教材、教員養成などが、後者は農村及び少数民族の

アクセスが挙げられる。また、質の向上分野でも、教員配置や学校マネジメントはまだあまり活動が行われていない。特定項目に複数の援助が行われている場合にも、同一目標達成に向けてそれらの関係性や補完性が明確なものもあれば、同じ項目への援助というのみでこれらが希薄なものも少なく、体系的はまだ十分には確保されていない。

これは上記援助の多くが、EFA-FTI導入前から実施されていたことに関係する。これらの援助は開始時期も様々であり、大半は2002年前後であるが、米のラジオ教育、独の農村ネットワーク支援のように90年代後半から続いているものもある。新たなプログラムの導入とこれに沿った援助の実施には通常タイム・ラグがある。本件の場合にも、EFA-FTIのもとで新規に開始された事業（米、学習標準等）もあるものの、多くのドナーはEFA-FTI導入を契機に事業を見直し、既存事業の協力期間終了にあわせてスクラップ・アンド・ビルドや新フェーズの協力を継続的に行う形でアラインメントを進めている。こうしたアラインメントの進展は、単一のプログラムに基づくPBAの付加価値と言えよう。

またホンジュラスでは、ハリケーン・ミッチ災害の復興支援を契機に現地ドナーにより活発に援助協調が図られてきており、教育分野でも1998年にMERECEが結成されている。MERECEでは毎月1回会合を開催のうえ、互いの活動の情報交換、EFA-FTIを含む教育分野の各種イシューの議論と共同取り組みなどが行われている。こうした中、JICAプロジェクトによって開発された算数教材がスウェーデン及びカナダの資金協力により全国配布用に印刷される、同じく算数分野の現職教員研修（5県をモデルに実施）がスペインによる現職教員研修支援を通じて全国展開に結びつく、などの協調が進展している。こうした援助協調もPBAの付加価値と考えられる。

しかしEFA-FTIの実施には、現在のプログラムのコンポーネント・項目のみでもまだ多くの介入が必要な状況にある。援助協調を図ったとしても、各ドナーのプロジェクト援助の積み上げのみで体系的な介入を確保するのは容易ではない。政府自身による介入がベースにあり、援助がこれを補う形で、全体として体系的な介入が可能となるよう

な体制の確保が必要となる。同ベースの構築の観点からは、政府の実施能力強化とコモン・ファンド支援の進展が、今後のEFA-FTIの計画実施の重要な鍵になっている。

(3) 実施プロセス

EFA-FTIは、実施面にも問題が見られる。現地でインタビューを行った11ドナーの全てから、教育省の実施能力の弱さと教員組合や政治の問題が、EFA-FTIの目標達成の制約要因として挙げられた。教育省の実施能力に関しては、EFA-FTI策定時のドナー・アセスメントでも強化の必要性が言及されている。しかしプログラムには特に反映されず、EA-FTI開始後に問題が顕在化し、2004年及び2005年の教育省・ドナー合同評価では同問題が主要議題となっている。コモン・ファンド設置に伴い急務となった予算管理能力及び調達能力強化については、コモン・ファンド・ドナーにより能力アセスメントと組織強化計画策定の支援が開始されている。また、教育省の能力強化を盛り込んだセクター・プランの策定も新たに検討されている。

PBAでは、対象国制度の活用が基本要素となっている。しかしコモン・ファンドでは、教育省の実施能力の弱さに伴う信用リスク管理の問題から、当初はホンジュラス側制度の活用ではなく世銀の制度に準じて実施するとされており、ホンジュラス側能力に比して格段に高い要求が課せられている。途上国のリーダーシップに関して、セクター・プログラムはホンジュラス側ではなく、ドナーの要求が起点となっている。

こうした要求への対応にホンジュラス側はかなりの時間・労力を費やしており、援助協調・調和化は図られているものの、ホンジュラス側の調整コストは必ずしも軽減されていない。

教育省の実施能力に関しては、ホンジュラスではEFA-FTIがはじめてのPBAということも斟酌する必要はある。しかし政権交代に伴う頻繁な人事異動や機構改革が、継続的な政策遂行や組織・人員強化の障害となり、実際に実施能力が弱いのも事実である。このためコンサルタントが教育省に配置され、援助事業を取り仕切ってきており、コ

ンサルタントと教育省職員との給与格差、コンサルタントから職員への知識・技術の移転の欠如なども問題となっている⁹。

教員組合は、ホンジュラスでは力が強く、給与や業務環境をめぐって政権との対立・ストが恒常化している。組合問題は、授業時間数の改善、人件費の削減、教員の業績評価、採用の適正化や配置転換の大きな障害になっている。政治は、既述のように政策の継続性の問題があり、例えば2006年の政権交代でも、新政権ではそれまでの教員給与削減策に代えて、給与のアップが打ち出されるといった状況にある。

この関連で、EFA-FTIのもと、2005年には教材全国配布が軍隊の協力も得て実施されたほか、コモン・ファンドにより貧困児童への学用品配布が行われている。しかし、これらも結果としては同年下期の大統領選挙を睨んだ政権（当時）による国民の人気取りに使われたとの指摘もある。特に学用品配布に関しては、当該コモン・ファンド・ドナーはホンジュラス側の主体的判断を尊重したとしているが、他ドナーからはEFA-FTIの優先分野の点で疑問視する声があった。

組合や政治といった問題は、一般的には「外部要因」と見なされる。また、教育セクター内だけでの対応は難しい。しかし組合問題の解決なくして、授業時間の適正化や教員配置の改善等は難しく、組合問題はEFA-FTIの大きなリスク要因となっている。EFA-FTI目標の実現のためには、教育セクターを超えたより上位（例えば大統領府）での対応も含め、同問題の改善や政策継続性の確保への取り組みを検討していく必要がある。

3. 事例2：ベトナム EFA-FTI¹⁷

ベトナムのEFA-FTIは、援助依存度の低い途上国の事例として、援助依存度と、プログラムへのアラインメントや体系的な介入、途上国の実施能力と途上国のリーダーシップ、政策の一貫性とプログラムの実施などについて示唆を示している。また、PBAの有効な国について考えるうえで興味深い事例となっている。

(1) 計画の適切性

ベトナムでは、1990年代を通じた初等教育普遍化への取り組みの結果、1990年から1998年で純就学率が86%から95%に、終了率も57%から74%に改善されてきた⁸。こうした中、2000年代の課題となっているのが、残す5%の就学率達成と、量のみでなく質の改善である。このうち前者については、男女格差は少ないものの、所得、民族、地域格差の問題があり、就学率では所得別5段階の最富裕層と最貧困層、人口の約9割を占めるキン族と少数民族、都市部と農村部の間では各々11%、12%、6%（2002年）の格差がある。

質に関しては、授業時間数の不足（年間総授業時間で国際標準の約3分の2）、その背景にある教室不足や教員給与問題による二部制の多さ（児童の約10%のみが全日制）、質の劣った校舎や教育設備、教員の質や教授法の問題が要因として挙げられる。なお、ベトナムでは近年の教員増員や学齢人口の減少もあり、教員は必ずしも不足していない。教員一人当たりの児童数も全国8地域で27-33人と国際基準（30人）のレベルにある。しかし国家基準の教員資格を持つ教員は1998年時点で約7割に過ぎない⁹。

こうした課題を背景に、2000年に策定された「社会経済開発戦略（SEDS）2001-2010」及び「教育開発戦略（EDS）2001-2010」、並びに各々の5カ年計画（2001-2005）では、初等教育の完全就学達成（純就学率99%）を目標として明示している。これに続く前期中等での完全就学レベル実現（純就学率74%から90%）を目標に加え、たうえで、僻地、少数民族地域、貧困層における教育強化、質の改善（カリキュラム、教員の質、教授法等）、行政システムの強化を戦略として掲げている。初等教育の完全就学達成と、質の改善、民族格差是正は、2002年に策定された「包括的貧困削減成長戦略文書」（CPRGS）にも盛り込まれている。

EFA-FTIは以上の計画・目標をベースに、EDSのもとでの新カリキュラム導入（2002年）や策定過程での追加的状況分析結果を盛り込んで策定されている。EFA-FTIは既存のSEDS、EDS等に代わるものではなく、2010年までを対象期間に異な

るフォーカスで策定された計画・目標を一貫した枠組みの中に関連付け、あわせ2015年までのより長期的視野を提供するものとされている（SRVN 2003）。これはベトナムでは、SEDSやEDSといった国家開発計画があくまで開発の基本文書であり、ドナーとの合意文書であるCPRGSやEFA-FTIはこれらを補完する文書と位置付けられていることによる。

EFA-FTIは、量から質への転換、初等・前期中等教育の普遍化、生涯学習の機会提供、コミュニティの参加動員、効果的なマネジメントと資源活用を戦略ゴールとし、幼児教育、初等教育、前期中等教育、ノンフォーマル教育を対象としてカバーしている。対象毎に「アクセス」、「質と適切性」、「マネジメント」の3領域における目標、2015年までのターゲット、アクション・プログラムを示している（表2参照）。

EFA-FTI全体としての包括的な対象・領域設定に加え、初等教育プログラムでは領域毎の目標実現に向けて、施設建設を含む量的拡大、ハード（施設）及びソフト（カリキュラム、教員、教材）の面での質の向上とその評価、配慮を要する多様な児童への質を伴った機会の提供と格差是正、中央及び地方レベルのマネジメント強化などの、幅広い取り組みが盛り込まれている。これはEFA-FTIが各種計画・取り組みを統合する形で策定されたことによる。ドナーの協力（UNESCO、世銀等の技術支援や、幅広いドナーとのコンサルテーション）や地方政府関係者の参画も得て、2年半をかけて準備されたこととも無関係ではない。

以上のようにEFA-FTIは、ベトナムの初等教育の課題及びEFA-FTIの目標に向けたプログラムとして体系性が高い。ベトナムでは地方分権化のもと国レベルでのEFA-FTIに基づき各省が省レベルでのEFA-FTIプログラム（省計画）を策定・実行する計画となっている。省計画は、64省中10省で計画手法の開発をかねてパイロット的に策定試行中の段階にある。EFA-FTIの成果達成は、いかに省計画が上位計画である国家EFA-FTIと政策体系としての一貫性を有し、かつ各省の課題に即した形で策定され、適切に実施されるかによってくる。

（2）介入の進展

ベトナムでは、政府予算¹⁰に占める教育セクターの割合は2004年時点で18%である¹¹。教育セクター予算の32%は初等、21%が中等前期に充てられている。教育セクター予算では、政府予算が73%、非政府予算が27%である。非政府予算には援助（ODA）のほか、教育債券、「教育の社会化」政策による保護者・地域・企業からの寄付等が含まれる¹²。ODAが教育セクター予算全体に占める割合は約1%に過ぎない。予算支出も、経常経費は75%で、25%が資本支出に充てられている。人件費は経常支出の71%であるが、教員給与は2001年時点で一人当たりGDPの1.45倍と、EFA-FTIの国際目標指標（3.5倍）を大きく下回っている¹³。ベトナムでは予算面における教育セクターの援助依存度は低く、2003-2015年通算でEFA-FTI予算の10%程度（2015年の計画値は1.6%）であり、ベトナム側が介入の主体となっている。

EFA-FTIへの2006年6月時点までの主な介入は表2のとおりである。各領域でアクション・プログラムに沿った介入が既に行われている。このうちベトナム側に関しては、EDSのもと、教育予算にイヤーマークの特別勘定を設け、初等教育普遍化及び識字撲滅の推進、教育プログラム及び教材改定、教員研修及び師範学校施設改善、山岳少数民族地域や僻地の教育支援、中核校、大学、職業訓練校などの施設改善、学校へのIT導入及び外国語学習推進、職業訓練能力強化の7つのプロジェクトから成る国家ターゲット・プログラム（NTP）を実施している。このうちからまでがEFA-FTIに関わる活動である¹⁴。また教育の質の改善に向けてアクティブ・ラーニング・アンド・ラーニング（ATL）への移行を盛り込んだカリキュラム改定が2002年に行われ、同適用に向けた取り組みが進められている。政府予算はじめ保護者・地域の寄付等の非政府予算財源をもとに、学校建設や施設改善も行われている。

EFA-FTIが5ヵ年計画（2001-2005）の途中で導入されたため、ベトナム政府は2005年までは実施中の主要活動を中心としてパイプラインをEFA-FTIにアラインさせる形で取り組みを進めきた。同成果をふまえて2006年からの5ヵ年計画期間にフォ

初等教育分野におけるプログラム・ベースド・アプローチの援助に関する評価分析
 ホンジュラス及びベトナムのEFA-FTIの事例から

表2 ベトナムEFA-FTIプログラムの概要と政府・ドナーの取り組み

目標	2015年までのターゲット	アクション・プラン	ドナー	支援内容
<p>【アクセス】</p> <p>1. 全児童、特に少数民族、困難な状況にある児童、女子への良質な初等教育アクセス</p> <p>2. 全児童による5年間の初等教育終了</p>	<p>・遠隔地における学校建設(2005年までに年間最低250校、2010年までに同100校)</p> <p>・留年減少(2010年までに2.5%、2015年までに1.0%)</p> <p>・退学率減少(2010年までに2.0%、2015年までに1.0%)</p> <p>・退学児童復学率(2010年までに前年退学児童の70%、2015年までに95%)</p>	<p>全児童への就学可能な学校施設提供</p> <p>全児童が初等教育5年間を終了するプログラム実施</p> <p>不利な状況にある児童や除外された児童に対するプログラム実施(学校整備、教員配置、教科書配布、一言語教育、英字釜、保護者啓発)に対する初等教育提供</p>	<p>政府、世界銀行、英国、EC、カナダ、ベルギー、ニュージーランド</p> <p>政府、世界銀行、英国、ルウエー、カナダ、オーストラリア</p> <p>政府</p> <p>小学校教育(政府予算、地域寄付金等)</p> <p>北部山岳地域小学校建設</p> <p>北部山岳地域コミュニティ学習センター普及</p> <p>イカレーション教育</p> <p>イカレーション教育</p>	<p>国家ターゲットプログラム(NTP)中、初等関連プログラムへの特定財政支援(TBS)I(関連NTP: 初等教育義務化・識字教育、山岳地域、少数民族、不利な状況の児童の教育支援)を5に不利な状況にある児童の教育支援、学校の質の基礎基盤達成に向けた施設改善、教材整備、行政・地域等能力強化(PEDC)</p> <p>小学校建設(政府予算、地域寄付金等)</p> <p>北部山岳地域小学校建設</p> <p>イカレーション教育</p> <p>イカレーション教育</p> <p>NTPへのTBS I(関連NTP: カリキュラム改革・教材整備、中核校施設整備、教員訓練・師範学校施設改善)</p> <p>NTP ITスタック訓練(学校へのIT導入)</p> <p>教員能力基盤確立、教員研修、行政能力強化支援(PTDP)</p> <p>新カリキュラム試行モデル、教員・校長・行政官研修</p> <p>北部山岳地域教員研修</p> <p>教授法、カリキュラム・教材開発、教員研修、機材供与</p> <p>施設改善(教育債券による施設強化プログラム)</p>
<p>【質及び適切性】</p> <p>3. 全学校での学校品質基準の確保など、質的拡大から良質な学習成果を伴った質の向上への転換</p>	<p>・全教員による年間30日の理職教員研修受講(2003年・)国家基準達成(2010年まで)</p> <p>・全教員への各学年特別科目指導要領の毎年配布</p> <p>・教員への大学院用英字釜、IT研修受講機会提供</p> <p>・経済成長、給与政策と整合性ある給与待遇</p> <p>・カリキュラム、教授法、教材の継続的改善と継続的評価システムの確立</p> <p>・教科書無料配布(2006年までに恵まれない地域の全児童、2015年までに全国の全児童)</p> <p>・教育支出増取(2015年までに児童一人当たりUS\$15 1校あたり4000US\$)</p> <p>・災害脆弱地域を優先に2010年までに全仮設校舎を固定構造に改善</p> <p>・全校で全日制導入(2015年まで)</p> <p>・年間授業時間増加(2015年までに国際標準の900時間)</p>	<p>新カリキュラム遂行(2002-07年)</p> <p>(教科書、教材配布、教員研修、IT)</p> <p>教員能力向上と研修実施(教員・校長研修、職業基盤開発、授業支援、質の保証制度確立等)</p> <p>児童の学業達成度評価(新方式試行、授業バックアップ開発、実施評価)</p> <p>学習環境と学習効果の質的改善(授業時間増加、教室新設・改善、施設基準導入、教科書配布)</p> <p>カリキュラム継続改善(2008-15年)</p> <p>(カリキュラム評価・改善、前期中等を含む8年間教育サイクルの準備)</p>	<p>政府、世界銀行、英国、EC、カナダ、ベルギー、ニュージーランド</p> <p>政府</p> <p>政府、世界銀行、英国</p> <p>日本</p> <p>ベルギー</p> <p>UNICEF</p> <p>政府</p>	<p>EMIS開発、地方レベルに適用、情報ベース政策策定能力強化(SOMET/SREM)</p> <p>国家EFAマネジメント・プラン、省レベルアクション・プランの計画策定支援(PREP/BCEP)</p> <p>国・省レベルの初等教育セクター開発計画策定及び策定能力強化支援(PEDP)</p> <p>中期支出計画枠組み(MTEF)策定支援</p> <p>私立学校設置・運営</p>
<p>【マネジメント】</p> <p>4. 中央、省、郡、学校レベルのマネジメント能力強化</p> <p>5. 地方分権マネジメント、9年生サイクル、量から質への転換を含む、包括的セクター開発・改革</p>	<p>・教育マネジメント情報システム(EMIS)、政府情報システム(GIS)、地方分権化が機能(2010年まで)</p> <p>・全国での質の高い私立学校運営</p>	<p>国レベルの情報ベース政策決定メカニズム開発、政策実施状況モニタリングシステム近代化、分権化マネジメント準備、量から質への転換枠組み、9年間マネジメント導入、私立学校開発、開発事業調整</p> <p>省、郡、学校レベルの計画運営能力強化(研修支援プログラム実施、予算管理システム改善、技術支援・品質管理ユニット設置)</p> <p>資源の有効活用、費用分担方式確立</p> <p>全レベルでの情報ベース政策決定メカニズム確立</p>	<p>政府、EC</p> <p>カナダ、UNESCO</p> <p>日本</p> <p>世界銀行</p> <p>政府</p>	<p>EMIS開発、地方レベルに適用、情報ベース政策策定能力強化(SOMET/SREM)</p> <p>国家EFAマネジメント・プラン、省レベルアクション・プランの計画策定支援(PREP/BCEP)</p> <p>国・省レベルの初等教育セクター開発計画策定及び策定能力強化支援(PEDP)</p> <p>中期支出計画枠組み(MTEF)策定支援</p> <p>私立学校設置・運営</p>
その他			<p>ルウエー</p> <p>世界銀行、ADB、カナダ、デンマーク、英国、EC、日本、オーストラリア</p>	<p>NTP全プログラム(7件)に対する支援(TBSとは別枠)</p> <p>貧困削減支援クレジット(PRSC JV)による教育セクター貢献分(PRSCの10%程度)</p>

(出所) JICA 2007、SRVN 2003、同2005a、2005b、World Bank ed. 2003、同2004、及びMinistry of Education DAD data をもとに筆者作成

ローアップを含むアクション・プログラムを加えている。

ドナー側の援助は、こうしたベトナム側の取り組みにアラインする形で行われている。その顕著な例がNTPに対する特定財政支援（TBS）であり、既述の7プロジェクトのうち から までを対象に2005年から128百万ドルの援助が供与されている。2004年のEFA-FTI導入後、比較的短期間で多額の財政支援が実施された背景には、CPRGSのもと、既に貧困削減支援借款（PRSC）による財政支援が行われていたことが挙げられる。

TBS以外にも、新カリキュラムにあわせた教員研修や基準の策定、山岳地域等における施設改善や行政・地域の能力強化などのプロジェクト援助が多数のドナーにより実施されている。世銀と複数ドナーの協調融資による初等教員開発プロジェクト（PTDP）や不利な状況にある児童に対する支援プロジェクト（PEDC）には、プロジェクトの予算総額の16-18%がベトナム側から投入されている。

実施中プロジェクトは、2001-2003年頃に開始された比較的新たな案件が多い。EDSを受けて始められたものもある。2000年以前には国連機関及び一部の二国間ドナー・NGOを除き初等教育分野の援助はほとんど行われておらず、2001年には策定が着手されたEFA-FTIが呼び水の役割を果たした面もあった。EFA-FTIがEDS策定直後のタイミングで一貫性をもって策定されたことも、EFA-FTI初期段階からのプログラムに沿った多数の介入の確保につながっている。

しかし、これらのプロジェクトはプロジェクト間の直接的な連携は図られていない。例えば世銀、日本、ベルギーそれぞれによる同じ教員研修分野のプロジェクトでは、互いの活動や経験・成果の共有はほとんどなされていない（JICA 2007）。援助協調に関しては、EFA-FTI導入に伴い2004年から教育セクター・グループが設置され、定期的な会合が開催されているが、活動内容は主に情報・意見交換である。これはベトナムが主導する形でEFA-FTIが進められていることが大きい。ドナー間のコミュニケーションは、情報共有や活動調整の不十分さが問題指摘されている（UNICEF 2004）。

（3）実施プロセス

ベトナムのEFA-FTIは、強いオーナーシップと低い援助依存度のもと、ベトナム側がリーダーシップをとって進められている。また、実施能力、実施環境も基本的には良好であり、プログラムの実施に影響を与えるような問題は生じていない。しかし、省レベルのEFA-FTIを担う地方政府の行政能力は、中央省庁に比べてまだ弱く、地域格差も大きい。開発成果達成の観点からは、分権化システムの中で省レベルにおける適切な取り組みをどのように確保していくか、省レベルでの取り組みに必要な予算の適切な配賦や技術支援をいかに進めていくか、が重要となる。

ベトナムでは、歳入はいったん国家予算に計上された後、各省に配分される。教育予算も、保護者・地域の寄付等は各省にそのまま割り当てられるが、他の財源は人口等を基準に経済社会状況などを考慮して配分される。省内の予算配分は省に全面的に委譲されている（瀬木ほか 2006）。しかし、省によって予算管理方法が異なり、中央と地方の予算管理方法が同時性を欠く、国レベルでの教育開発目標や方向性と予算配分が必ずしも整合しない、地方、中央、セクター・レベルにおける予算執行状況の適宜・的確な把握やとりまとめが難しい、といった課題が生じている。イヤーマーク予算であるNTPに関しても、多くの実績は挙げているものの、達成すべきターゲットの一層の明確化と結果に関する定期的な報告が必要との指摘がされている（SRVN 2005c）。これらは今後の省レベルにおけるEFA-FTIの実施においても、計画・実施方法のある程度の標準化、国レベルEFA-FTIと省レベルのEFA-FTIの一貫性の確保、進捗に関するモニタリングや報告制度の確立等が課題であることを示している。

4. 比較・考察

ホンジュラス及びベトナムは、初等教育普遍化に向けた単一プログラムのもと、政府・ドナーがこれにコミットして取り組みが進められており、政府においては資金計画にもとづく予算割り当て

が、ドナーにおいては援助協調を通じたプログラムの支援が行われているとの共通点を有する。また、財政支援を含む調和化など、同じような取り組みが行われている。他方、同じEFA-FTIでもプログラムのつくりは異なっており、同プログラムの実施状況や、その結果としての開発成果達成に向けた進展には違いがある。

(1) プログラム策定

ホンジュラスとベトナムのプログラムの相違をまず見てみたい。ベトナムのEFA-FTIは初等教育の普遍化と質の向上に広く取り組む内容である。これに対してホンジュラスのEFA-FTIは、フォーカスが絞られた限定的なプログラムである。フォーカスされた分野の目標達成に必要な取り組みも必ずしもカバーされていない。次に、ベトナムのEFA-FTIは既存の戦略との整合性がとれているのに対し、ホンジュラスのEFA-FTIは既存戦略との不整合が見られる。

なぜこのような相違が起こったのか。第1に、EFA-FTI策定プロセスを見ると、ベトナムではドナーの支援を得つつ時間をかけてプログラムが策定されたのに対し、ホンジュラスの場合には短期間にドナーの参画も不十分な形でプログラムが策定されている。第2に、この背景としてホンジュラスでは、対象国選定プロセスに間に合わせるために、至急プロポーザルを提出しなければならないというタイム・プレッシャーがあった。その意味では、(おそらくはEFA-FTI導入年の対象国承認という実績確保のために)短期間の作業を課したEFA-FTI事務局の進め方にも問題があった。第3に、ホンジュラスは援助依存度が高く、独自の開発事業展開には制約があった。よって開発プログラム策定にあたっては、既存の援助介入や今後の援助の可能性を考慮せざるを得ない面があった。特にプログラムの実効性が求められるEFA-FTIの場合には、この点が顕著であった。このため目的からではなく手段を重視してプログラムを策定することになった。結果として、目的に基づき策定する場合に比べてプログラムの体系性の確保には制約が生じた。援助規模が限られているので、特にその傾向が大きかった。

プログラムへの取り組みが開発成果達成につながるには、計画が適切である必要がある。計画の適切性に関し、ホンジュラスとベトナムのEFA-FTIの以上のような相違からは、2つの留意点が導かれる。

第1が、対象国の援助依存度の考慮の重要性である。PBAでは、途上国の主体的プログラムがベースとなるが、援助が一部の補完的介入でしかない国と、援助依存度が高く、プログラムの計画内容自体が援助によって左右されるような国とでは置かれた事情が異なる。後者のような国の場合、開発成果実現にドナーが負う共同責任はより大きい。途上国の課題にあった体系的なプログラムが、援助の中期的展望も得たうえで策定されるためには、プログラム策定段階から途上国・ドナー間の十分な対話が図られることが必要である。PBAは一律に議論されがちであるが、こうした国による違いをふまえたうえで、対象国の状況に合った適切なパートナーシップのあり方を検討することが重要である。

第2が、国際イニシアチブの現地化にあたっての、対象国の既存の開発戦略との整合性に対する配慮の必要性である。ホンジュラスでは、既存の開発戦略との整合性が十分でなく、限定的なEFA-FTIの導入により、似ているが少しずつ違う内容の開発戦略の並立と政策体系の混乱が生じている。また、国際合意に基づく開発戦略の方が重視される中、限定的なEFA-FTIがより広い取り組みを含む政府開発戦略にとってかわった状況を生んでいる。

ベトナムでは、自国の政策体系をあくまでベースとしてEFA-FTIと政府開発戦略との関係を整理するとともに、EFA-FTIが5ヵ年計画の半ばで策定されたことへの対応を、次の5ヵ年計画にあわせてとることで自国開発戦略とEFA-FTIとの一貫性の確保が図られた。しかしこれは、非常に強いオーナーシップを有し、援助依存度も低く、マネジメント能力を有する国であればこそ可能であったとも言える。また、ベトナムの場合も、5ヵ年計画の半ばでさらに上位の10ヵ年計画を超える期間を含むEFA-FTIの策定・導入には相応の工夫を要している。

PBAのもとで国際合意に基づく新たなプログラ

ムを開始する場合には、対象国の既存の開発戦略との整合性、政策体系としての一貫性、またプログラム策定・導入のタイミングを十分考慮すべきと言えます。特に国際イニシアチブの現地化にあたっては、援助依存に起因する途上国側とドナーとの交渉力の違いにも配慮し、押し付けや政策の二重体系にならないような、ドナー側の思慮が求められる。

(2) プログラム実施

次にプログラムへの介入の進展と実施体制を見てみよう。体系的な介入の確保、対象国実施能力、安定した実施環境、ステークホルダーとの関係構築などがプログラムの実施、ひいては開発成果の実現可能性に影響を与えている。

ベトナムでは同国自身の取り組みをベースに、これにアラインする形で援助が行われている。実施能力が高いことから早い段階で多額の財政支援が行われ、新規援助が増加する中、プログラムを広くカバーする形で体系的な介入が短期間に進んでいる。「教育の社会化」のもと、保護者、地域、企業といった広範なステークホルダーとの協同関係が構築されていることも、プログラムの介入に寄与している。政策・組織の継続性が確保されていることが、EFA-FTIへの取り組みをESDの5ヵ年計画のローリングに反映させることができる理由となっている。

他方、ホンジュラスでは、援助が介入の中心、かつ新規援助よりも主に既存の援助のスクラップ・アンド・ビルドでプログラムへのアラインメントが進められており、介入の体系的性はまだ十分には確保されていない。援助もドナーが直接実施するプロジェクトがほとんどであり、介入のメイン・アクターはドナーである。したがってPBAについても実態はドナーが主体となつての援助協調との状況にある。また、実施能力の問題から、財政支援は供与までに時間がかかり、供与額が限られている。信用リスクの恐れから、途上国制度ではなくドナーの制度を適用し、ホンジュラスの能力以上の高い水準の要求が課されている。政権交代に伴う政策・組織・人事の変更、教員組合問題が実施プロセスの阻害要因になっている。

こうしたベトナムとホンジュラスの相違からは、開発成果実現に向けたプログラムの実施に関して以下の留意点が抽出できる。

第1が、実施環境の吟味の重要性である。政治、経済、社会情勢は、プロジェクトにも影響はあるが、プログラムの場合には上位のセクターあるいはマクロ・レベルの開発政策とのつながりがより直接的であり、対象も包括的であるため影響は一層大きい。取り組みも長期にわたるため、政策の継続性や安定した取り組みを可能とする実施環境がより重要となる。政策の継続性等は「外部要因」として位置づけられがちであるが、むしろ「前提条件」とも言える。PBAにおける対象国のコミットメントは安定した実施環境への政治意志なども含めて考える必要がある。

同様に、プロジェクト・レベルでは「外部要因」でも、プログラムの場合には内部要因として取り組むべきものもある。その例がホンジュラスの教員組合問題であり、教育セクターの主要利害関係者（組合）に関わる問題解決と開発成果実現に向けた協同関係の構築はステークホルダー・リレーションである。開発成果実現に向けたプログラムの適切な実施のためには、広い視点から利害関係者を捉え、協同関係を構築することが求められる。

第2に、対象国の実施能力の考慮の必要性である。実施能力は、先行研究でもPBAの進捗に影響を与える要因として挙げられている。しかし問題は、実施能力の制約でPBAが進まないことではなく、実施能力をふまえた進め方がされていないことである。その例がホンジュラスであり、ファンド設置の後に問題対応として実施能力強化が行われている。また、既述のように信用リスク管理のためにドナーの制度が適用されている。このような場合には、対象国の負担の面でも、実施能力を強化したうえでファンド設置等を図るような段階的進めの方が適切と考えられる。PBAについては一律な議論や進め方がなされる傾向が見られるが、PBAの基本要素は手段であって目的ではない。各国固有の状況をふまえたうえで、開発成果実現に向けた途上国主導の取り組みのために何が必要かに基づいて、適切なエントリー・ポイントとルートで進めることが重要である。

(3) 援助の貢献の評価

以上では、開発成果実現の観点から、PBAのあり方について見てきた。では、こうしたPBAにおいて、開発成果の実現に対する援助の貢献、とりわけ特定ドナーの貢献はどのように評価すべきなのか。

国際援助社会では、特定国の就学率や終了率の向上といったレベルの開発成果の達成には対象国及び関係ドナー全体による包括的な取り組みが必要との認識のもと、SWApにはじまり現在のPBAに至るアプローチを進めてきている。そうした取り組みにより達成された開発成果と援助の関係については、特定ドナーの介入との厳密な因果関係(attribution)ではなく、同介入が上記取り組みにどのように貢献したかの可能性(plausibility)の観点から評価する考え方が広く適用されるようになってきている。貢献の可能性の評価手法にはまだ確立されたものはないが、一般的には取り組み全体の中での戦略的な位置付け、他のアクターとの協調、当該ドナーの介入における直接的なアウトカムの達成度等を総合的に勘案して評価が行われている(三輪 2005)。同様の評価は、JICAでも既に実施している(JICA 2006、2007)。

本論では、特定ドナーの援助から対象範囲を広げて、取り組み全体のベースとなる対象国のプログラム、同プログラムへの介入全体及び開発成果の実現の可能性を評価した。EFAに向けた課題解決のためには、教育の質の向上一つをとっても教員研修、カリキュラム、教材、学校施設、教育行政、学業達成度の評価などの総合的な取り組みが必要である。事例からも、開発成果と特定ドナーの介入の関係性は貢献に基づく考え方で捉える必要があることが示された。

また、特定ドナーの援助のみでなくプログラム全体を鳥瞰し、異なる国の事例を比較することで次のような視点も抽出された。

第1は、プログラムの開発成果実現の中でどのような介入の貢献がより大きいかということである。当然ながら援助は各ドナーの経験や比較優位を活かす形で行われ、これ自体には妥当性がある。また、プログラムが適切なものであれば、そのコンポーネントに対する介入は開発成果に貢献する

介入と言える。プログラム全体の成果を考えると、介入の波及効果は同じではない。政策・制度や行政機関の実施能力がサービス・デリバリーを含むセクターの機能の制約となっているような場合には、プログラムの適切な実施の基盤となる政策・制度の支援や行政機関の能力強化は、教材配布、教員研修、学校建設といった個別活動への介入よりも波及効果が大きい。投入が1人のアドバイザーであっても、規模の大きい学校建設よりもプログラム全体への貢献が大きい場合もある。

第2は、同じような介入でも、国によって意味が異なるとの点である。上記のようなケースとは異なり、対象国が実施能力を有し、資金やハードの不足が問題であるような場合には、求められるのは技術協力よりも、資金供与や学校建設・整備である。

JICAはホンジュラスとベトナムでEFA-FTIに対して協力を行っているが、ともに現職教員研修、かつ特定地域を選んでの試験的事業であっても、意味が異なる、プログラムの中での位置付けが違うからである。ホンジュラスでは、算数分野の教材開発・教員研修を行っているが、算数分野はJICAがリード・ドナーであり、教材は全国の教員に活用されている。このため上記教員研修は、特定地域での試行に留まらず、全国に展開されていくべき位置付けにある。他方、ベトナムでは、ATLを盛り込んだ新カリキュラムの普及にあわせ、「児童中心型教授法」に関する教員研修が行われている。ベトナムでは、別途、世銀等がATLのための教員研修をより広域的に展開しており、同協力では教員の能力基準の策定も組み合わせられている。北部山岳地帯ではベルギーによる教員研修も実施されている。これら3つの取り組みのいずれが今後の全国的な教員研修の核となっていくかを考えると、基準づくりも盛り込み、多くの省で実施されている世銀等の研修が可能性として高い。JICAの協力では他省への普及が上位目標に掲げられているが、上記現状に照らせば、広域展開を図るための試験事業というよりは、世銀等の支援に基づく研修と組み合わせられていくべき補完的アプローチ(「児童中心型教授法」)を提示するものと位置付けられる。日本の初等教育援助で多い現職教員研修については、ほとんどが普及を

目指す試行取り組みとの位置付けである。しかし、個々の国の状況や他ドナーによる同様の援助との関係によって、同取り組みの持つ意味や達成すべき成果は異なる。位置付けを明確にして協力を行っていくことが必要である。

5. おわりに

本研究では、手段に焦点を当てる形で一律に議論されがちであるPBAについて、開発成果の達成にどのようにつながっていくのかとの視点から分析することで、各国の状況に応じて進め方や手段の選択を吟味する必要性を明らかにした。PBAは、途上国のリーダーシップの重視、途上国の単一・包括的プログラムと予算枠組みに基づく援助、援助協調・調和化、途上国制度の活用を要素とするが、これらの要素を開発成果に結びつけるためには、計画の適切性、体系的な介入の進展、実施プロセスの3点を丹念に検証することが重要である。特に開発成果の実現に向けたより効果的なPBAの実施に関する議論を深めるためには、援助依存度、対象途上国の実施能力、実施環境の3つを重視することの必要性を、本論事例の検討結果から強調したい。

この関連で、先行研究では、PBAは援助依存度の高い国において特に有効であるとされている。援助の乱立、重複等への対応としての援助協調・調和化等の進展、すなわち援助の投入の仕方の視点で見れば、援助依存度の高い国でPBAが有効なのは当然である。他方、本研究のようにPBAが開発成果の実現につながるかの視点で見た場合、援助が多く援助調整を要する国の方が有効なのは言うまでもないが、援助依存度が高い国で特に有効とは簡単に言えない。一定適度の主体的介入が可能で援助依存度と実施能力を有する国の方が、途上国のリーダーシップや開発成果によりつながっていることが、ベトナムとホンジュラスの事例では示されているからである。ただし、PBAがどのような国で特に有効かは今後のさらなる研究課題である。

最後に、以上を踏まえ、我が国のODA評価に対する示唆をしておきたい。我が国のODA評価

は、途上国の開発戦略との合致の有無及び個々の投入やアウトプットをもって評価が行われてきた傾向がある。また、開発成果よりも日本の援助の効果に主として関心が払われ、「帰属」の観点から効果を評価しようとする傾向がある。結果として、日本の援助もその一部である全体的な取り組みや、その中での日本の援助の具体的な位置付けについて必ずしも十分に吟味されてきていない。このような傾向は、現在、我が国ODAにおいて進められつつある、プログラム・アプローチにも引き継がれている。開発成果実現に向けた途上国プログラムへのアラインメントというよりは、日本の援助をより効果的・効率的に組み合わせるための「日本のプログラム」の発想が主となっているからである。しかし、「日本プログラム」のみでは、基本的に開発成果の達成が難しいのは既に記したとおりである。開発成果実現につながるより戦略的・効果的な援助の実施のためには、対象国のプログラムの体系、これに対する途上国、日本、他ドナーの介入、その中での日本の援助の位置付けなどを、より広く捉えて援助を計画し、その効果を評価していくことが肝要である。

付記

本稿の見解は、筆者個人のものであり、JICAの統一的な見解を示すものではない。最後に、貴重なご指摘をいただいた匿名の差読者に感謝申し上げます。

注記

- 1 複数国の事例研究と総合分析を組み合わせた主要研究として、初等教育分野ではNetherlands Ministry of Foreign Affairs 2003、教育分野を含むPBAではBrown 2001、Foster 2000などが挙げられる。
- 2 ホンジュラスは、9-3-4制の教育制度を採用しており、各3年の3サイクルからなる9年間を基礎教育としている。初等教育は最初の6年間2サイクルにあたる。EFA-FTIの対象は最初の2サイクルであることから、本稿では、初等教育の用語を使用する。
- 3 2005年にドナー会合に提出された案では、教育の効率性と質、インフラ整備、組織強化（関係組織

- の組織開発と地方分権、人材の能力開発)が3本柱。セクター・プログラムではあるが各段階(中等、高等教育を含む)の詳細の計画は言及されておらず、EFA-FTIの補完といった色彩が強い。
- 4 以下、政府予算データは財務省資料(SEFIN 2006)及び現地調査時教育省説明資料(Presentacion del Presupuesto de la Secretaria de Educacion 2005)による。
- 5 EFA-FTIのインデカティブ・フレームワークの目標指標では、教員給与は一人当たりGDP3.5倍程度、教育関係経常費における教員給与以外の経費33%。ホンジュラスでは前者6倍、後者52%となっている。
- 6 EFA-FTI実施の一元的な所掌部署として2003年に教育省に設置されたEFA局も政治的任命の局長のもと、主要業務はコンサルタントによって行われてきた。しかし、教育省の本来組織・業務へのEFA-FTIの統合や多数のコンサルタントに要する経費が問題となり、教育省・ドナー合同評価の結果、2005年末に廃止された。
- 7 ベトナムのEFA-FTIは就学前教育、初等教育、前期中等教育、ノンフォーマル教育の4分野をターゲットとするが、ここでは初等教育分野を対象に見る。ベトナムの教育制度は5-4-4制の教育制度である。
- 8 教育統計については、SRVN 2003、SERVN 2005b、UNESCO 2000に基づく。
- 9 以前は前期中等終了後、教員養成校で3年コース履修(9+3)であったが、後期中等教育終了後、教員養成校で2年コース(12+2)にグレード・アップされた。ただし資格を持たないものや以前の資格のみの者も少なくない。
- 10 以下、予算データはSRVN 2005bに基づく。
- 11 EFA-FTIでは2015年までに同割合を20%にすることが実施指標の一つに掲げられており、2004年時点で計画に1年先んじる形で18%が達成されている。
- 12 ベトナムでは、教育財源の拡大・多様化に向けて2003年から教育債権を発行している。社会全体で教育を支える「教育の社会化」政策も1990年代から進められており、保護者、地域、企業から寄付や積立金を募り、学校建設・補修、各種教育新興活動が行われている。
- 13 教員給与については、教育の質の向上の観点から最低賃金の引き上げ等が政策として打ち出されている。EFA-FTIでも同政策に沿った改善が実施指
- 標に掲げられているが、同達成後も2015年で一人当たりGDPの2.19倍程度の水準に留まる(SRVN 2003)。
- 14 NTPには、2001-2005年の間、毎年教育予算支出の3-4%が支出されてきている(SRVN 2005a)。

参考文献

- 国際協力機構(2006)『特定テーマ評価・プログラム評価(ホンジュラス基礎教育分野)』
- 国際協力機構(2007)『特定テーマ評価・プログラム評価(マラウイ、ベトナム教育分野)』
- 国際協力事業団(2002)『開発課題に対する効果的アプローチ:基礎教育』
- 潮木守一ほか(2006)『初等教育の普遍化戦略に関する事例研究:平成14-17年度科学研究費補助金基盤研究(B)(1)研究成果報告書』
- 三輪徳子(2005)「国別プログラム評価の困難性を越えて-実践的方法論に関する考察と試案」、『日本評価研究』5(1):27-44.
- Bartholomew, Ann, Leurs, Robert and McCarty, Adam (2006). *Joint Evaluation of General Budget Support 1994-2004: Vietnam Country Report*, OECD/DAC.
- Brown, Adrienne, Foster, Mick, Norton, Andy and Naschold, Felix (2001). *The Status of Sector Wide Approches, Working Paper 142*, ODI.
- CIDA (2003). *CIDA Primer on Program-Based Approaches*.
- CIDA (2004). *Honduras Project Approval Package: Project No. A-32111-001*.
- European Commission (EC)(2003). *Guidelines for EC Support to Sector Programmes*.
- Foster, Mick (2000). *New Approaches to Development Co-operation: What can we learn from experience with implementing Sector Wide Approaches? Working Paper 140*, ODI.
- FTI Development Committee (2006). *Progress Report for the Education for All-First-Track Initiative*.
- Honduran Ministry of Education (SE)(2002). *Fast Truck Initiative Education for All Honduras 2003-2015 FTI-EFA*.
- Mesa Redonda de Cooperantes Externos en Educadion (MERECE)(2002). *Evaluation de la Propuesta del*

- Gobierno de Honduras para participar en EFA-FTI.*
 Netherlands Ministry of Foreign Affairs (2003). *Local Solutions to Global Challenges: Towards Effective Partnership in Basic Education.*
- OECD/DAC (2003). 「評価と援助の有効性：評価と結果重視マネジメントにおける基本用語集」
- OECD/DAC (2005). *DAC Guideline and Reference Series Harmonizing Donor Practices for Effective Aid Delivery vol. 2.*
- Secretaria de Educacion, Republica de Honduras (SE) (2004a). *Memoria de la Pevision Conjunta del Plan EFA.*
- Secretaria de Educacion, Republica de Honduras (SE) (2004b). *Informe Pevision Conjunta del Plan EFA 2002-2004.*
- Secretaria de Educacion, Republica de Honduras (SE) (2005a). *Memoria de la Pevision Conjunta del Plan EFA.*
- Secretaria de Educacion, Republica de Honduras (SE) (2005b). *Plan Estrategico Sectoral de Educacion Periodo 2005-2015.*
- Secretaria de Finanzas, Republica de Honduras (2006). *Informe de la Liquidacion del Presupuesto General de Ingresos y Egresos de la Republica: Ejercicio Fiscal 2005 Gobierno Central.*
- Socialist Republic of Vietnam (SRVN) (2001). *The Education Development Strategic Plan for 2001-2010.*
- Socialist Republic of Vietnam (SRVN) (2003). *National Education for All (EFA) Action Plan 2003-2015.*
- Socialist Republic of Vietnam (SRVN) (2005a). *5 Year Strategic Education Development Plan 2006-2010.*
- Socialist Republic of Vietnam (SRVN) (2005b). *Vietnam Managing Public Expenditure for Poverty Reduction and Growth: Public Expenditure Review and Integrated Fiduciary Assessment.*
- Socialist Republic of Vietnam (SRVN) (2005c). *ODA Strategic Framework for Period 2006-2010.*
- UNESCO (2000). *Education for All 2000 Assessment*
- UNESCO (2005). *Education for All: The Quality Imperative, EFA Global Monitoring Report 2005-.*
- UNICEF (2004). *Potential for Education Sector Wide Approaches (SWAps) in East Asia, Proceeding Document.*
- Walford, Veronica (2003). *Defining and Evaluating SWAps, A Paper for the Inter-Agency Group on SWAps and Development Cooperation, Institute for Health Sector Development.*
- World Bank ed. (2003). *Vietnam: Partnership for Development, An Informal report prepared for the CG Meeting for Vietnam, Hanoi, December 2003.*
- World Bank ed. (2004). *Moving Towards 2010; Vietnam Partnership Report 2004, An Informal report prepared for the CG Meeting for Vietnam, Hanoi, December 2004.*

(2007.3.2受理)

An Evaluation Study of Program-Based Approaches (PBAs) for Development Assistance in the Primary Education Sector: A Case Study of the Education for All-Fast Track Initiative (EFA-FTI) in Honduras and Vietnam

Satoko Miwa

Japan International Cooperation Agency (JICA)

Miwa.satoko@jica.go.jp

Abstract

Program-based approaches (PBAs) are increasingly being applied to development assistance in the primary education sector. To examine their current operations and challenges, the author conducted a case study of the Education for All-Fast Track Initiative (EFA-FTI) in Honduras and Vietnam carried out under PBAs, focusing on progress towards development results. The study shows that, despite similar approaches and tools employed under the EFA-FTI, the program and progress between the countries differ due to factors such as aid-dependency, capacity, and an enabling environment. This paper argues that differentiated approaches based upon specific country contexts are essential. The importance of capacity development of partner countries is also emphasized. The author raises the issue of evaluating aid under PBAs; she proposes that evaluation be based upon the concept of contribution and points out the importance of assessing strategic positioning and the role of aid, given overall efforts towards the attainment of development results.

Keywords

program, development results, aid effectiveness,
contribution, capacity development

【研究論文】

段階評価における項目の重み導入による三次元的分析の試み - 大学における授業評価分析を事例として -

星野 敦子 北原 俊一 新行内 康慈 安達 一寿

十文字学園女子大学 社会情報学部

hoshinoa@jumonji-u.ac.jp

綿井 雅康
十文字学園女子大学
人間生活学部

牟田 博光
東京工業大学大学院
社会理工学研究科

要 約

評価結果を政策に生かす場合、評価項目の重要度の違いがその後の政策を左右する要因となる。すなわち、重要度の高い項目の評価の方が低い項目に比較して、改善された場合に総合評価にもたらす影響が大きくなる。本研究は、大学における授業評価を事例として、評価項目の段階評価に対し、項目の重みを導入した三次元的分析を行うことで、授業評価の結果を効果的に授業改善に結び付ける方法を検討した。項目の重み付け法としてはコンジョイント分析を用いた。各要因の部分効用値をX軸にとり、各要因と関連する授業評価項目の段階評価データをY軸にとって、要因の相対重要度をバブルの大きさとするバブルグラフを作成することにより、評価項目の重みと段階評価を可視的に表すことが可能となった。本研究によって提案された方法は、より広い評価分野への応用への適用が望まれる。

キーワード

コンジョイント分析、授業評価、データ解析、アンケート調査、大学教育

1. 背景

近年、わが国でも多くの大学において学生による授業評価が実施されている。アメリカにおいて、学生が教育評価に関わるようになったのは1950年代からであり、このころから学生による評価に関する文献も見られるようになる(Falchikov 2005)。1960年当時、すでにアメリカの大学・カレッジの40%で授業評価が行われていたとされている

(Stecklein 1960)。

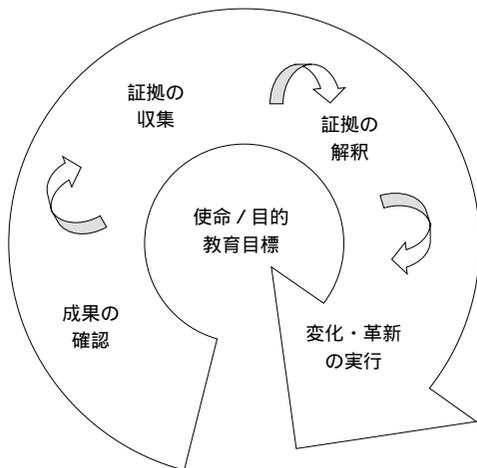
一方わが国においては、平成3年の大学設置基準の改定により、自己点検・自己評価・ファカルティ・ディベラップメント(FD)の重要性が示されたことが、授業評価の実施を促進した(示村1992)といわれている。文部科学省の調査によれば、平成10年度に学生による授業評価を実施したのは、国公立大学全体で334大学(約47%)に過ぎなかったが、その後増加の一途をたどってい

る。平成15年度で授業評価を実施している大学は、国立96大学（約99%）、公立68大学（約89%）、私立469大学（約99%）であり、大学全体では633大学（約91%）となっている（文部科学省高等教育局大学振興課 2005）。

Braskamp & Ory（1994）によれば、学生によるシステム化された授業評価の増加は 高等教育機関における質の評価の総体的増加 深刻な財政問題に直面した経営者の意思決定のための材料の必要性 学生による授業の段階評価は、効率的に収集でき、「教育の質」という難解な概念を定量化できる、といった要因の結果である。

アメリカにおいては学生による授業評価がしばしば教員評価（査定）の材料として利用されてきたが、現在主流となっているのは、評価結果を改善に活用するものである。Maki（2004）が提案している「学習評価のサイクル」（図1）は、教育観に基づいて、教育目標・目的を定める。教授学習活動についての成果（アウトカムズ）を確認するための証拠を収集し、その解釈を行い、その結果から、革新を行い、再び評価サイクルを繰り返していくものである。この評価サイクルはいわゆる「アウトカムズ評価」であるが、評価が査定として終わっているのではなく、改善に向けて繰り返している点に特徴がある。

図1 学習評価のサイクル



（出所）Maki.P.L.（2005）

一方わが国においては、授業評価が人事に関わる教育業績の査定に直接用いられることはほとんどない（大山 2001）。授業評価の主要な目的は、教育効果の向上のための授業改善である。文部科学省の報告書でも、「学生による授業評価」は「授業の質を高めるための具体的な取り組み」として位置づけられている（文部科学省高等教育局大学振興課 2005）。授業評価が広く普及した現在、多くの大学の興味は、授業評価の結果をいかにして授業改善に生かすかという点にある。

2. 研究の目的

評価においては、その評価結果を改善に生かすことが求められる。しかるに、評価の低かった項目すべてを一度に改善することは不可能である。そこで項目の中でも、より重要度が高く、改善により大きな成果が期待できるものから対応していくことが求められる。

本研究では、評価項目の重要度（重み）を測定し、これを段階評価（rating）と組み合わせることによって、有効な改善策を3次元的に提示する方法を実証的に示すことを目的としている。本研究で提案する方法は、授業評価のような質的な評価分野よりも、むしろ政策評価等に用いる方が的確であるが、ここでは著者の数年来の研究対象としての授業評価を対象とした分析を行う。

重み付けの方法は大きく2つに分類できる。1つは、回答者が直接重要度を示す方法であり、Grading法（最も重要な要因を10点満点として点数付けをする）や順位法（すべての要因の順位または順位に相当する点数をつける）がこれにあたる。もう1つは回答者の選好に関するアンケートに基づいて、要因の重要度を求める方法であり、AHP（Analysis Hierarchy Process：階層分析法）やコンジョイント分析がある。

評価者が学生であることから、直接要因の重要度を求めても、有効なデータは期待できない。またAHPは一対比較を繰り返す方法であることから、授業時間の一部を割いて行う方法としては不向きである。そこで、本研究ではコンジョイント分析を用いて評価項目の重み付けをすることとし

た。

コンジョイント分析は、複数の評価要因のもつ水準（属性）に対する回答者の選好データから、各要因の重要度と好まれる属性を求めるものである。初めLuce and Tukey（1964）により、計量心理学分野において開発されたが、その後マーケティングリサーチの分野で広く利用されるようになった。Green and Wind（1973）は、マーケティングにおける多属的意思決定のためのモデルについて初めて示したとされている。1990年代に入ってから、環境経済学の分野でも研究が開始された（栗山 2000）。また近年では、いまだ事例は少ないものの、教育ならびに福祉の分野にも応用されるようになった（真城 1998: 長井 2002など）。

コンジョイント分析を授業評価に用いる場合、最も重要なのは各要因の水準の設定である。学生は実際に自分が受けた授業を基準としてしか水準の是非を判断できないため、水準は実際の授業を基準に設定し、また各水準の示す内容について、調査票に明解に示す必要がある。

3. 研究の方法

コンジョイント分析を大学における学生による授業評価に利用したデータがないことから、本調査に先立ち、異なる3大学で実施された4種の授業（講義科目2科目、演習科目2科目）を対象としたプレ調査を実施した。調査結果から、コンジョイント分析の授業ニーズに対する有効性が確認された。学生は単に「楽な」授業をよしとするわけではなく、自分にとってより高いパフォーマンスを得られる授業を求めることが、成績や専攻の違い

による分析を通して明らかとなった。（星野・牟田 2003: 星野 2004）。本調査では調査は、授業の段階評価とプロフィールによるコンジョイント分析の両方を実施した。

分析の対象となったのは、A女子大学において平成14年度後期に開講されたコンピュータリテラシー科目（演習科目）である。1年生の必修科目であり、習熟度別3クラス留学生クラスを加えた4クラス編成であるが、このうち習熟度別3クラスのみを分析対象とした。段階評価アンケートの有効サンプル数は、上級クラス37、中級クラス41、初級クラス38、コンジョイント分析調査の有効サンプル数は、上級クラス40、中級クラス44、初級クラス42である。コンジョイント分析では上記科目について授業の条件（要因）とその水準（属性）を設定した。表1は設定した授業の要因と水準を表している。要因は6つで、それぞれが2つの水準を持つ。

調査は複数の要因を同時に示す「全概念法」を採用し、各水準については、以下の説明を加えて調査票を作成している。

知識伝達の量：シラバスで示した内容はカバーするが、必要最小限の知識量でよいか、あるいはできるだけ多くの知識を得たいか

難易度：基本的な内容を学習すれば十分か、ある程度高度な内容も含めたほうがよいか

教材：板書を中心とした講義がよいか、プレゼンテーションなどで前もって構成された教材を利用した講義がよいか

課題：授業の中で実施する小テストや小レポートなどの課題を実施したほうがよいか、そのような課題は必要ないか

テキスト：テキストを利用したほうがよい

表1 分析のための要因と水準

要因 水準	知識伝達の量	難易度	教材	課題	テキスト	個別指導
1	少ない	易しい	板書中心	課題なし	なし	なし
2	多い	難しい	プレゼン中心	課題あり	あり	あり

（出所）著者作成

か、テキストは必要ないか

個別指導：授業中に教室を回って必要に応じて個別指導をした方がよいか、必要ないか

以上の6つの要因（各2水準）について直交計画を生成し、各水準を組み合わせた8種類のプロファイルを抽出した。さらにモデルの適切さをチェックするためのホールドアウトカードを2種類に加え、計10種類のプロファイルについてアンケートを行った。ホールドアウトカードは部分効用値の推定には用いず、推定された部分効用値に基づいて算出されたホールドアウトカードの評価値を回答者の実際の評価値と比較する。両者の相関が高ければ推定結果は信頼できるといえる。質問形式は「完全プロファイル評定型」とし、すべてのプロファイルについて、「ぜひこの条件で受講したい(5)」から「この条件では受講したくない(1)」までの5段階で評価してもらった。各要因の水準がすべて名義尺度であることから、「Discrete（離散）モデル」を適用し、効用値の推定は最小二乗法を用いた。

一方、段階評価アンケートについては、授業評価に関する15の評価項目に加え、学生の授業に対する期待・努力の度合い、理解度、出席率、および総合的な満足度の計20項目を設定した。出席率および満足度は10%から100%の間で評価してもらい、その他の項目については5段階評定法を利用している。

コンジョイント分析の6要因と授業評価項目を関連付け、コンジョイント分析の部分効用値と相対重要度、及び関連項目の授業評価データをバブルグラフ上で表すことにより、要因の重要度と項目の評価の連携を図った。

4. 結果と考察

(1) 各要因の部分効用値と相対重要度

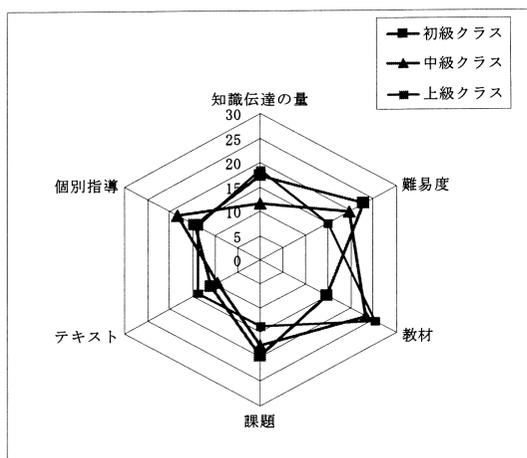
はじめにコンジョイント分析の結果について検討する。表2はクラスごとの部分効用値ならびに

表2 部分効用値と相対重要度（クラス別）

要因	水準	初級クラス (N=42)		中級クラス (N=44)		上級クラス (N=40)	
		部分効用値	相対重要度	部分効用値	相対重要度	部分効用値	相対重要度
知識伝達の量	少ない	-0.039	17.16	-0.003	11.60	-0.084	18.37
	多い	0.039		0.003		0.084	
難易度	易しい	0.313	22.95	0.264	19.70	0.022	14.83
	難しい	-0.313		-0.264		-0.022	
教材	板書中心	-0.176	14.69	-0.338	23.25	-0.359	25.39
	プレゼン中心	0.176		0.338		0.359	
課題	課題なし	0.182	19.75	0.122	17.58	0.122	13.77
	課題あり	-0.182		-0.122		-0.122	
テキスト	なし	0.051	11.01	-0.037	9.45	0.016	13.69
	あり	-0.051		0.037		-0.016	
個別指導	なし	-0.158	14.44	-0.287	18.42	-0.116	13.96
	あり	0.158		0.287		0.116	
CONSTANT		2.878		2.889		2.966	
Pearson's R		0.993		0.999		0.996	

(出所) 著者作成

図2 クラスごとの相対重要度



(出所) 著者作成

相対重要度を、また図2はクラスごとの相対重要度のグラフを示している。

効用値が正の値を示している水準が好まれており、その値が大きいほど要因の重要度が高い。相対重要度は部分効用値の全体における割合を示している。なお、完全プロフィール評定型においては個人の推定値が算出されることから、表2では、効用値・相対重要度共に平均値を示している。したがって表2においては部分効用値の割合は相対重要度とは一致しない。

表2を見ると、クラスにより重視する要因が異なり、初級クラスでは「難易度」、中・上級クラスでは「教材」の相対重要度が最も高くなっている。各要因に関して好まれる水準は、3クラスとも「知識伝達の量」が「少ない」、「難易度」は「易しい」、「教材」は「プレゼン中心」、「課題」は「なし」、「個別指導」は「あり」となっており、「テキスト」については中級クラスのみが「あり」、初級・上級クラスは「なし」となっている。

要因のうち、「知識伝達の量」については各クラスとも平均相対重要度に比較して部分効用値が小さい(初級クラス:0.039、17.16% 中級クラス:0.003、11.60% 上級クラス:0.084、18.37%)。このことから、分散が大きく、「多い」を好む学生と「少ない」を好む学生とで意見が分かれていると考えることができる。また「テキスト」につ

いては部分効用値も相対重要度も比較的小さいことから、「あり」でも「なし」でもあまり影響がないことがわかる。すなわち、テキスト自体を再検討すべきであるということになる。

「難易度」は初級・中級クラスで相対重要度も部分効用値も高いことから、授業内容が「易しい」ことが重要であるということになる。

一方、上級クラスでは部分効用値が0.0219と小さく相対重要度も14.8%とあまり高くない。これは上級クラスの学生にとって授業の難易度があまり重要ではないことと、「易しい」と難しい」を好む学生で意見が分かれていることを示している。

図2からも明らかなように、難易度は習熟度クラスレベルが低いほど相対重要度が高く、「理解できる授業」を切望している様子が伺われる。「個別指導」に関しては、特に中級クラスにおいて相対重要度が高い。これは、中級クラスの学生は他のクラスと比較して習熟度に幅があり、その差に対応する必要があるためであると考えられる。

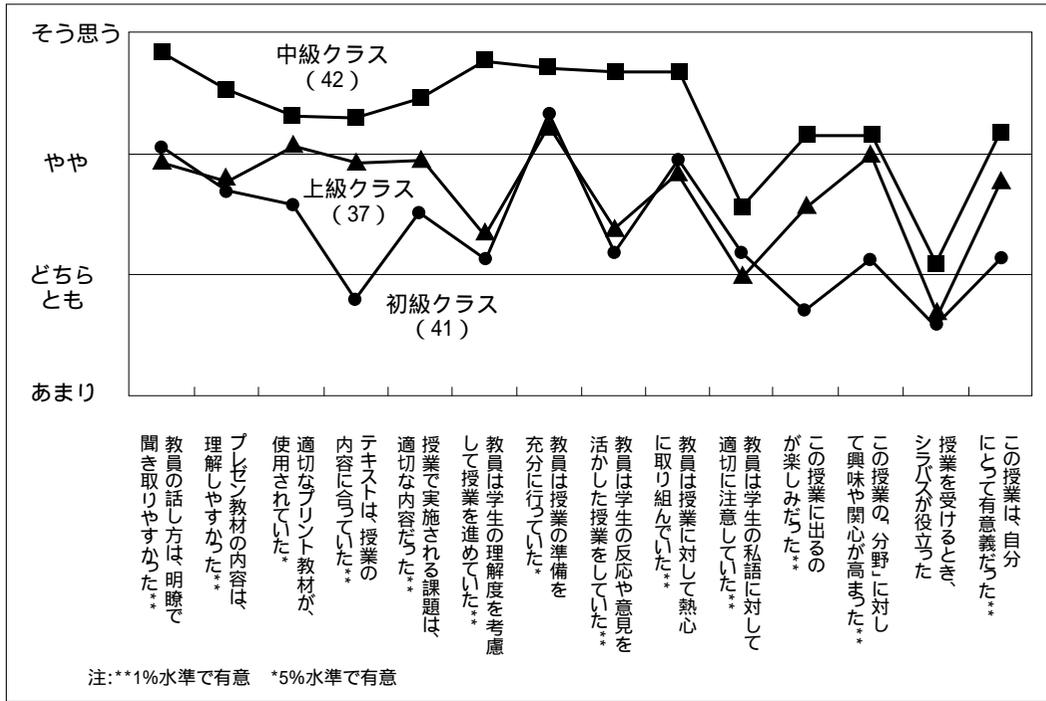
(2) 段階評価の概要

評価項目は、授業評価に関する15項目に加え、学生の授業に対する期待・努力の度合い、理解度、出席率、および総合的な満足度の計20項目である。図3は評価項目のうち、「非該当」の多かった項目を除いた14項目のクラス平均値ならびに分散分析の結果を示している。

図3を見ると、「授業を受けるときシラバスが役立った」以外はすべてクラスによる有意差が認められた。総体的に見て、中級クラスの評価が高く、初級クラスは低くなっている。

項目別に見ると、教員側の努力である「話し方」「授業準備」「熱意」等の評価が高い。学生とのコミュニケーションに関する項目である「理解度考慮」「学生の反応」及び「私語」などはクラスによる差が大きい。また教材に関連した項目である「プレゼン教材」「テキスト」および「課題」についてもクラスにより評価が分かれている。

図3 授業評価項目の平均値（クラス別）



(出所) 著者作成

表3 コンジョイント分析の要因と授業評価項目

コンジョイント分析		学生による授業評価（段階評価）
要因	水準	評価項目
知識伝達の量	多い 少ない	この授業の「分野」に対して興味や関心が高まった
難易度	易しい 難しい	教員は学生の理解度を考慮して授業を進めていた
教材	板書 プレゼン	プレゼン教材の内容は、理解しやすかった
課題	あり なし	授業で実施される課題は、適切な内容だった
テキスト	あり なし	テキストは、授業の内容に合っていた
個別指導	あり なし	教員は学生の反応や意見を活かした授業をしていた

(出所) 著者作成

(3) 関連評価項目の設定

授業評価データとコンジョイント分析とを連携した分析を行うために、双方の関連項目について検討を行う。表3は連携項目を示している。コンジョイント分析の要因と対応評価項目については、以下の観点から連携させた。

- ・「課題」「テキスト」について、はいずれも「あり」「なし」の2水準であるので、対応する評価項目としては「授業で実施される課題は適切な内容だった」及び「テキストは授業の内容に合っていた」を関連項目とした。
- ・「教材」については、「板書中心」「プレゼン中心」の2水準を取っており、3クラスとも「プレゼン中心」を好んでいることから、「プレゼン教材の内容は理解しやすかった」を関連項目とした。
- ・「知識伝達の量」については、「少ない(必要最小限の知識)」と「多い(できるだけ多くの知識)」の2水準を取っている。関連項目としては、その分野に対する興味の度合いが知識量と関連が深いと考えられることから、「この授業の分野に対して興味や関心が高まった」を関連項目とした。

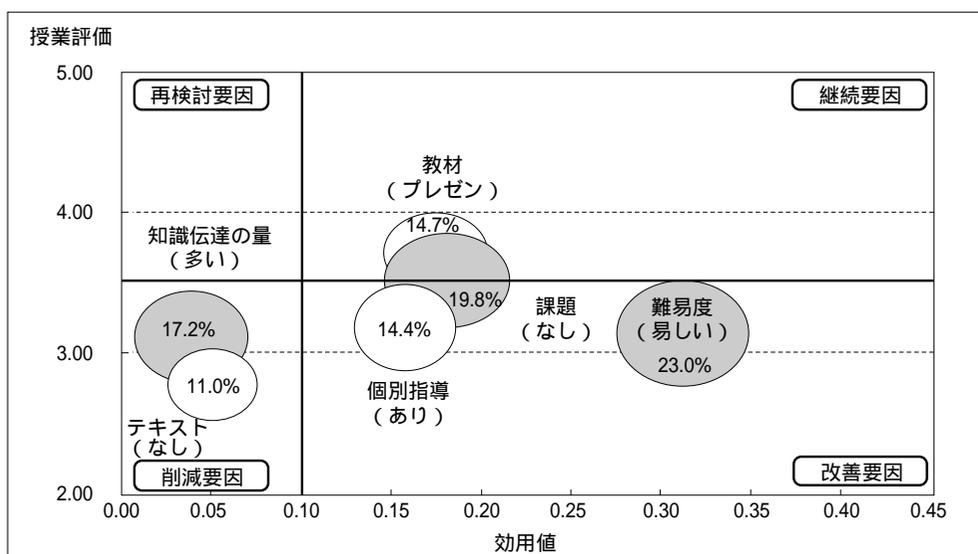
- ・「難易度」は「易しい(基本的な内容)」と「難しい(ある程度高度な内容)」の2水準をとっており、学生の理解度と関係の深い要因であることから「教員は学生の理解度を考慮して授業を進めていた」を関連項目とした。
- ・「個別指導」については学生個人への対応という観点から、「教員は学生の反応や意見を活かした授業をしていた」を関連項目とした。

(4) 段階評価に対する重みの導入

コンジョイント分析の各要因と授業評価項目を連携させ、段階評価に対して各要因の重みを導入したとき、各要因の部分効用値・相対重要度と授業評価の間どのような関係が認められるであろうか。ここでは3者の相互関係を明らかにするために、三次元による視覚化が可能となるバブルグラフを利用する。図4、図5、図6は各クラスのグラフを示している。X軸に部分効用値、Y軸に授業評価をとり、バブルの大きさは各要因の相対重要度を表している。網掛けのバブルは相対重要度が15%以上のものである。

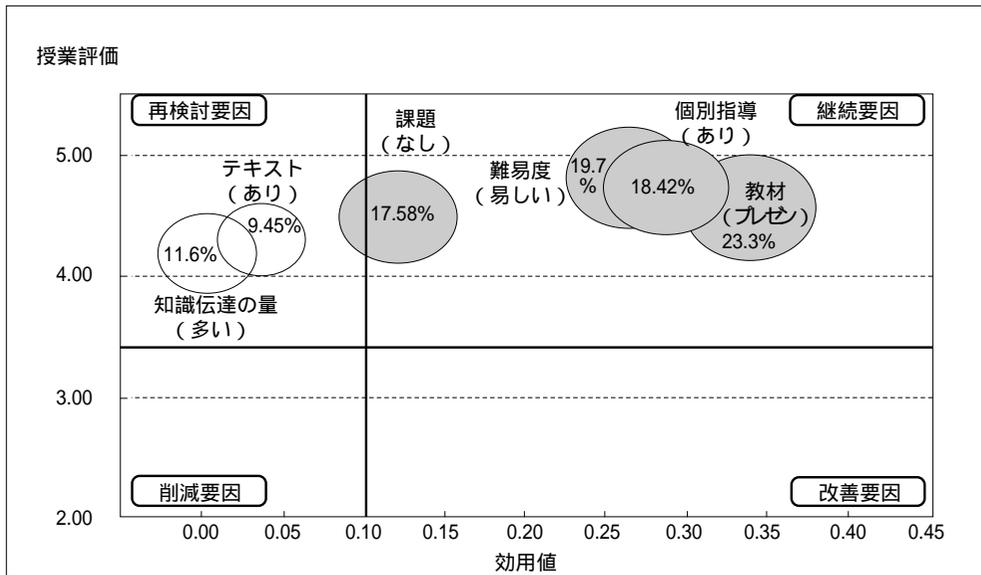
さらに、バブルの大きさと位置により要因の特徴を明確化するために、効用値(X軸)と段階評

図4 要因の重みを加えた授業評価と効用値(初級クラス)



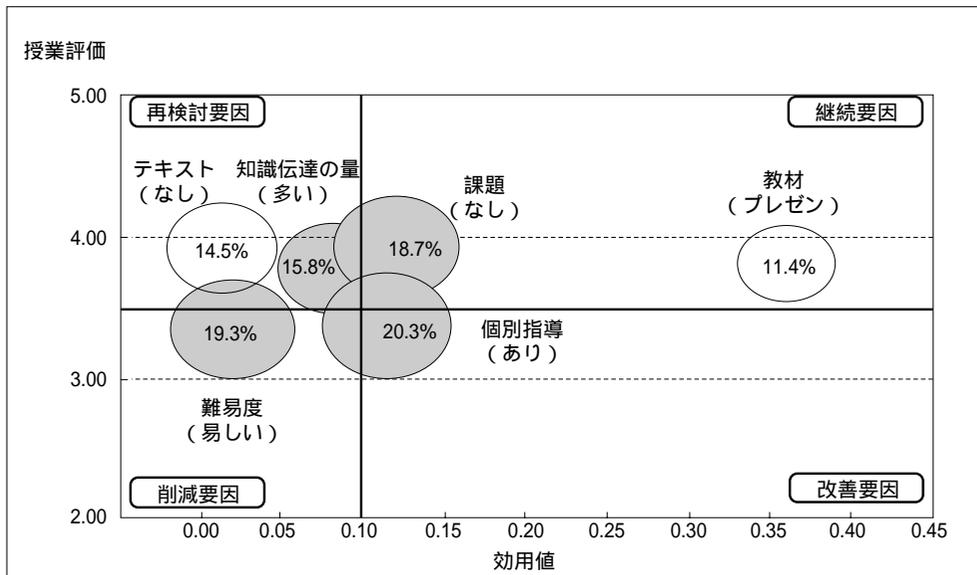
(出所) 著者作成

図5 要因の重みを加えた授業評価と効用値（中級クラス）



(出所) 著者作成

図6 要因の重みを加えた授業評価と効用値（上級クラス）



(出所) 著者作成

価（Y軸）をそれぞれ境界線で区切り、グラフを4象限に分割した。段階評価については3が「どちらともいえない」であるが、学生は全体的にやや高めの回答をするので、3.5を境界とした。また効用値についてはプレテストの結果から6要因のうち効用値が0.1以上となっているのが半分の3要因である場合がほとんどであることから、0.1を境界とした。各象限に含まれる要因の特徴は以下の通りである。

継続要因：効用値が大きく、授業評価も高い要因は、現状のままで満足度が高いものと考えられるため、この象限に入る要因を「継続要因」とする。すなわち、現状のまま継続していくべき要因である。

改善要因：効用値は大きい授業評価が低い要因は、授業全体の満足度に対する影響力が高いにもかかわらず、評価が低い要素であり、改善が必要であることから、「改善要因」とする。

再検討要因：効用値は低い、授業評価が高い要因は、教授努力が現れているにもかかわらず、授業全体の満足度に及ぼす影響が小さいことから、要因自体のあり方を再検討すべきであると考えられる。そこでこれらを「再検討要因」とする。

削減要因：効用値が低く、授業評価も低い要因は教授努力により改善しても、結果的に授業の満足度に及ぼす影響は小さく、努力が無駄になる可能性が高い。そこで、削減できるものは削減し、そうでないものはそのまま保留すべき要因であると考えられる。したがってこの象限に入るものを「削減要因」とする。

ただし、「再検討要因」と「削減要因」については相対重要度についても考慮する必要がある。部分効用値が小さく、かつ相対重要度が高い場合には部分効用値の分散が大きく、好みの違いの個人差が大きいことを表している。すなわち、部分効用値も相対重要度もグループの平均値であるために、部分効用値の好みの違いが大きいと結果的

に平均値は0に近づき絶対値が小さくなるのである。一方、相対重要度に関しては、グループ平均値が大きいということは、その要因を重要であるとする学生が多いということになる。したがって、相対重要度が高いのに部分効用値が低い場合には、学生はその要因を重視しており、好みがかかっているということになるのである。

このように考えると、「削減要因」と分類されるが相対重要度が高いものについてはむしろ「改善要因」であるといった方が的確である。ところが好みの水準が明確でないためにどの方向に改善したらよいかは明らかとなっていないため、改善方針を定めるのは容易ではない。また「再検討要因」として分類され、かつ相対重要度が高いものについては、教授努力の成果は現れているが、学生の好みがかかっているということになり、好みの違いはあっても評価が高いことから、むしろ「継続要因」であると考えてよい。

図4は初級クラスの授業評価と効用値のバブルグラフである。「難易度」及び「個別指導」が「改善要因」であることから、「学生の理解度を考慮した授業」ならびに「学生の反応や理解度を生かした授業」をするように授業改善を行う必要がある。「教材」についてはボーダーではあるがプレゼンのまま継続するが、やや内容を改善すべきであると考えられる。「課題」については現状「あり」であるが、学生は「なし」を好んでいる。しかしながら「継続要因」と「改善要因」のボーダーであることから、現状のまま内容の改善を行うのが適当であると思われる。「知識伝達の量」と「テキスト」は「削減要因」となっている。「知識伝達の量」については相対重要度がやや大きいことから学生の好みの多様性に対応する形で改善を行っていくのが望ましい。また「テキスト」はなくてもいいのではないかとと思われる。

図5は中級クラスの授業評価と効用値のバブルグラフである。「教材」「個別指導」「難易度」は「継続要因」であり、従来どおりの指導方法でよいと思われる。また「課題」についても学生は「なし」を好んでいるが、「継続要因」であることから、現状（あり）のままで十分である。「テキスト」と「知識伝達の量」についてはともに「再検討要因」となっている。「テキスト」について

はテキストそのものを見直す必要があり、別のものを利用するか、なしのままの体制をとるかを決定すべきであろう。また「知識伝達の量」については変化させたとしても総合的な満足度に影響を与えるとは考えにくいので、今のところは保留でよいと考えられる。

図6は上級クラスの授業評価と効用値のバブルグラフである。「教材」と「課題」は継続要因となっており、現状のままよいと思われる。また「テキスト」は「再検討要因」であることから、中級クラス同様見直しが必要である。「知識伝達の量」も「再検討要因」であるが、「継続要因」に近く、相対重要度も高いことから、現状維持でよいと思われる。

「個別指導」は「改善要因」であるが、相対重要度が高い割に「削減要因」に近いことから、学生による好みの個人差も大きいと考えられる。したがってよりいっそうの「学生の反応や意見を活

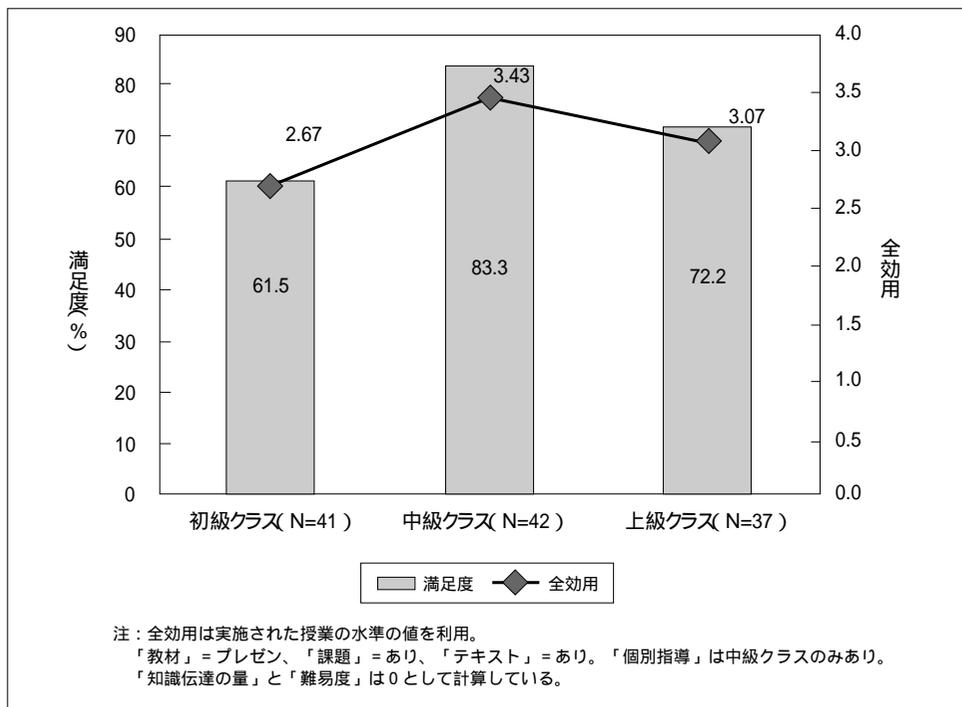
かした授業」への努力が求められる。「難易度」は「削減要因」であるが、やはり相対重要度が高いことから「改善要因」とみなすことができる。学生の好みが分かれているので、理解度に幅があるという前提での指導が必要となる。

(5) 授業満足度と全効用の検討

学生による授業評価においては、授業満足度の高さが、総合的な授業の良さを決める重要な要因となる。そこで、満足度に関する調査結果と、コンジョイント分析の結果について比較を行った。コンジョイント分析における総合的な満足度は、各要因の特定の水準における全効用の推定値(部分効用値の合計に定数項を加えたもの)で表すことができる。

この場合の水準は、実際に行われた授業に合わせる。図7はクラスごとの授業の満足度とコンジ

図7 授業満足度と全効用



ョイント分析の全効用の推定値をグラフに表したものである。

各要因の水準のうち、3クラス共通しているのは、「教材（プレゼン中心）」「課題（あり）」及び「テキスト（あり）」の3要因である。また「知識伝達の量」と「難易度」は基準が明確でないために、いずれも部分効用値を0として計算している。「個別指導」は中級クラスのみが「あり」で初級・上級クラスは「なし」である。図7から明らかのように、授業評価アンケートによる授業の満足度の大きさと、コンジョイント分析による全効用の推定値の大きさはほぼ同様の傾向を示している。授業満足度は初級クラス61.5%、中級クラス83.3%、上級クラス72.2%であり、分散分析の結果も1%水準で有意差が認められた。一方全効用の推定値は初級クラス2.67、中級クラス3.43、上級クラス3.07となった。

授業満足度は、授業に対する満足度をパーセンテージで答えてもらったものであるが、各クラスとも授業の要因ごとに条件を合わせて求めた全効用と同程度の平均満足度を示していることから、学生は直観的かつ総合的に満足度を判断しており、その結果は信頼性が高いとみなすことができる。なお、本研究の分析対象となった授業の授業満足度と他の評価要因の関係については、Hoshino et al. (2005) において詳しい分析を行っている。

5. まとめと今後の課題

本研究では、従来の授業の段階評価について、コンジョイント分析を用いて項目の重み付けを行い、評価結果を改善策に生かす方法について検討した。その結果、バブルグラフを用いて、段階評価、部分効用値、および相対重要度を3次的に示すことによって可視的に要因（評価項目）の特徴を示すことができた。さらにグラフを4象限に分割し、各象限に含まれる要因の特徴により、「継続要因」「改善要因」「再検討要因」および「削減要因」と類型化することで、グラフにおける要因の位置とバブルの大きさにより、迅速に要因の特徴が示されることが明らかとなった。

分析対象となった授業評価データは、事前にコンジョイント分析との連携を考慮していたわけではない。したがって要因の関連項目は必ずしも適切であるとはいえない。今後は両者の連携を前提とした、改善に結びつく調査項目・要因の設定による分析が必要であり、さらに各要因の水準の客観性を高めることが望まれる。また、完全プロファイル評定型においては個人の推定値が算出されることから、授業評価においても個人が特定されるデータが得られるならば、さらに事前の知識、成績、ならびに出席率といった他の要因の影響についても分析が可能であると思われる。

本研究において示された分析方法は、授業評価のような質的な評価よりも、むしろ政策評価等に適していると考えられる。本研究の実験としての成果を踏まえて、他分野においても分析を試みることで、三次元的評価分析手法をより実践的なものとしていくこと望まれる。

参考文献

- Braskanp, A. & Ory, J. C. (1994). *Assessing Faculty Work*. The Jossey-Bass.
- Falchicov, N. (2005). *Improving assessment through student involvement*. Routledg Falmer.
- Green, P.E. and Wind, Y. (1973). *Multiattribute Dicions in Marketing: A Measurement Approach*, Hindsdale . Dryden Press.
- 星野敦子 (2004)「大学における授業満足度シミュレーションの試み」、『日本評価学会春季第1回全国大会発表要旨集録』、175-180
- Hoshino, A., Kitahara, S., Shingyouchi, K., Adachi, K., and Watai, M. (2005). Analysis of Influential Factors on Computer Literacy Course Evaluation. *SITE 16th International Conference*, 872-877.
- 星野敦子・牟田博光 (2003)「コンジョイント分析を利用した大学における授業改善の試み」、『日本教育工学会第19回全国大会講演論文集』、691-692
- 栗山浩一 (2000)「コンジョイント分析」、大野栄治編著『環境経済評価の実務』、勁草書房、105-122
- Luce, R.D. and Tukey J.W. (1964). Simultaneous Conjoint Measurement: A New Type of Fundamental Measurement. *Journal of Mathematical Psychology*, 1, 1-27.

Maki, P.L. (2004). *Assessing for Learning*. stylus-AAHE
文部科学省高等教育局大学振興課 (2005)「大学における教育内容等の改革状況について2」文部科学省
長井克己 (2002)「コンジョイント分析による高専学生の英語に関するニーズ調査」、『津山工業高等専門学校紀要』、44 : 123-126
大山泰宏 (2001)「大学教育の展望と課題」、『京都大学高等教育研究』、7 : 37-55
真城知己 (1998)「大阪府下における特殊教育書学校

教員の大学における教員免許状取得への要望に関する調査」、『発達障害研究』、20 (3) : 91-97
示村悦次朗 (1992)「大学教育と授業評価 - 大学審議会の考え方」、『IDE現代の高等教育』、332 : 14-17
Stecklein, J. E. (1960). Colleges and universities programs: Evaluation. in Harris, C.S.(ed.) *Encyclopedia of Educational Research*. Macmillan: 285-289.

(2007.1.30受理)

Three-dimensional Analysis Incorporating the Weight of Factors for Evaluation – A Case for the University Course Evaluation –

Atsuko Hoshino Shun-ichi Kitahara Kouji Shingyouchi Kazuhisa Adachi

Faculty of Social Informatics, Jumonji University
hoshinoa@jumonji-u.ac.jp

Masayasu Watai

Faculty of Human Life Science
Jumonji University

Hiomistu Muta

Tokyo Institute of Technology
Graduate School of Decision Science and Technology

Abstract

Recently, “ student ratings ” of courses are becoming popular among Japanese colleges and universities in getting more information for Faculty Development (FD). The purpose of this paper is to propose a new three-dimensional analysis incorporating the weight of factors in student evaluations of university courses.

It is useful for the policy of FD to consider the importance of factors. If we know which factors are the most important for students, we can improve our course performance more effectively.

Conjoint analysis was used to determine the important criteria. We translate not only the results of conjoint analysis (“ utility ” and “ averaged importance ”) but “ students ratings ” data in a three-dimensional bubble graph. This graph makes it possible to show faculties the priority of factors that must be improved in their courses. Four divisions of the graph, named “ continue-factors ”, “ reexamine-factors ”, “ improve-factors ” and “ reduce-factors ” represent the features of the variables that are located in each division.

Keywords

conjoint analysis, course evaluation, data analysis, questionnaire, higher education

【研究論文】

Economic Burden of Disease and Cost Benefit Analysis of Emergency Medical Service (EMS) System in Romania

<p>Takao Ozaki The University of Tokyo minerva@sky.sannet.ne.jp</p>	<p>Marius Enescu The Ministry of Health, Romania enescum@hotmail.com</p>	<p>Moazzam Ali The University of Tokyo moazzam@u-tokyo.ac.jp</p>	<p>Chushi Kuroiwa The University of Tokyo ckuroiw@m.u-tokyo.ac.jp</p>
--	---	---	--

Abstract

The objective of the research is to provide an analytical model of the economic Cost-Benefit Analysis (CBA) and resultant policy recommendations, with the proposed emergency medical service (EMS) project in Constanta County (714,923 population, 2005), Romania, in view. The economic cost of investment is US\$ 31.5 million as per 2006 price level. The economic benefits are due to (i) incremental benefits attributable to EMS at the Constanta Emergency Hospital (CEH) as measured by willingness to pay (WTP) for healthier life, and (ii) social cost-saved (recovery of Productivity foregone due to emergent injury and diseases) measured by disability-adjusted life years (DALY). WTP was estimated by Contingent Valuation Method (CVM)-Double-bounded Dichotomous Choice method (449 interviewees) resulting in US\$ 1.6 million par annum (US\$ 214.8 per household, 8.9 percent of average household disposable income in Constanta). DALY in Constanta in 2005 is estimated at 118,885 person-years while applying the "ratio method" of total DALY to death DALY (7,286 deceased) and the disability adjustment factors (weighed average 0.67). With his and the productive beneficiaries (age group 15-59) of EMS at CEH (5,812 DALY) and the annual household disposable income in Constanta average (US\$2,420.7), financial-benefits-converted to the social cost-saved is US\$ 12.7 million (SCF=0.9), totaling to US\$ 14.3 million of economic benefit per year. The economic internal rate of return (EIRR) is 32.2 percent, thereby quantitatively revealing high economic viability (efficiency in scarce resource allocation) of the concerned EMS project in Constanta.

Keywords

Economic Cost Benefit Analysis (E-CBA), DALY, Willingness to Pay (WTP),
Economic Internal Rate of Return (EIRR), Emergency Medical Service (EMS), Romania

1. Introduction

In 1997, Romania embarked on the health sector reform shifting from the centralized government financing and delivery of services to a more

decentralized and pluralistic approach. In this light, the government introduced a compulsory health insurance fund, with 19 percent of gross wages and pensions being deducted. Yet the sector is facing a number of policy issues to be addressed, inclusive of the

mismatch between demand and supply of qualified and enough emergency medical services (EMS) in local cities in particular. As pointed out by the World Bank, the policy issues now the Romanian health system confronts include, among others, the followings.

- Shortcomings in the efficiency, equity and transparency of sector financing,
- Inefficient use of physical capacity and human resources in health care delivery,
- Weaknesses in governance of the system and the legislative framework;
- Critical infrastructure deficiencies resulting from inadequate maintenance and investment;
- and
- Clients' dissatisfaction with and distrust on health services delivered.

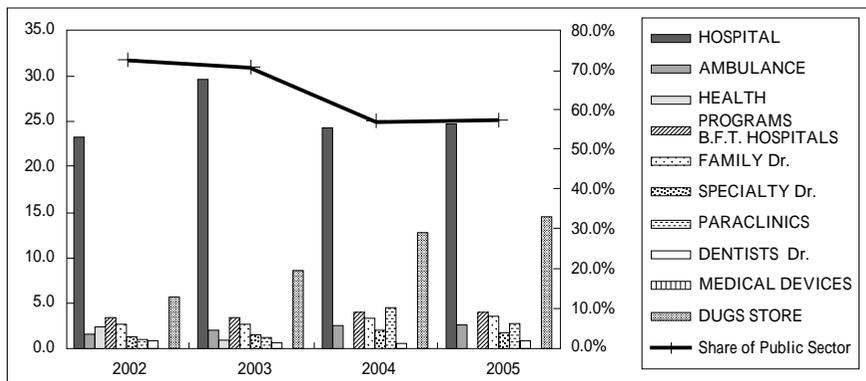
In Constanta with the fourth largest population in the country, the Constanta Emergency Hospital (CEH) is the largest general hospital in Constanta County with 66,008 in- and out-patients (2005), accounting for 58.7 percent of the aggregate patients of 112,486¹. Of this, emergency service beneficiaries accounts for 33.5 percent in average. The objective of the proposed public intervention measure (the project) is to provide more accessible services with increased

quality and quantity in a bid to maximize the health sector outcomes and impact for those requiring emergency medical services in the CEH, with around one-third of CEH patients. Specifically, the project is to (i) supply medical equipment urgently in need to the Emergency Department, (ii) train medical staff for the increased quality of patient care and survival rate of patients arriving at emergency room (ER), and (iii) upgrade the existing communication system of the County Directorate of Public Medical Pre-hospital Emergency Service enabling voice and data communication between the county ambulance dispatcher and sub-stations /ambulances. As depicted below (Fig.1), the share of public sector medical services in the Constanta County has recently been declined from above 70% in 2002 to less than 60% in 2005, with an increase in the use of private drug stores. This may reflect people's skepticism on the public health system and its service quality.

2. Objective and Rationale

The international financing institutions (World Bank and others) as well as academic institutes (Murray and Lopez, 1996, and Peabody J.W. et al. 2005) has since last decade endeavored to quantitatively analyze and estimate the value of life

Figure 1 Transitional Change in the Use of the Public and Private Sector Med-Services in Constanta



Source: Casa Judeteana de Asiguran de Sanatate, 2006

and healthy life in a logically consistent framework for analysis, while elaborating in tandem (proxy) measurement indices including the willingness-to-pay (WTP) and disability-adjusted life years (DALY). In this light, the objective of the research is to provide an analytical model of the economic Cost-Benefit Analysis (CBA) and resultant policy recommendations, while undertaking the quantitative analysis of the proposed emergency medical service (EMS) project in Constanta County, Romania. This research work is most likely the very first academic venture applied to Romania and other transition economies with a rigid framework of cost-benefit analysis (CBA) and the economic viability index (EIRR) based on the methodological research works on site. In tandem, the output of this task is used for substantial feedback, as a policy analysis and recommendations, to the government and policy makers in the light of efficient allocation of scarce resources in the country.

3. Methodology

3.1 CBA and EIRR analysis

The economic internal rate of return (EIRR)² analysis takes place herewith where economic viability of the Project is quantitatively measured by EIRR and the economic net present value (ENPV)³, based on the

estimated economic cost and benefit streams laid down over the project period. Financial costs are converted to economic costs in real term to reflect the true value of goods and services employed during the project period. Economic benefits are quantifiably valued in monetary terms by WTP for health and DALY.

3.1.1 Direct and indirect benefits

Health sector projects would generate a diverse set of benefits inclusive of productivity gained due to a fewer life years lost to mortality and morbidity, consumption gains due to a higher quality of life and increased life expectancy, and cost savings on curative treatments. Nonetheless, in line with what has commonly been in place, the economic benefits in this research is strictly confined to the productivity impact, thereby treating only the production side of the health effects, while eradicating the consumption side.

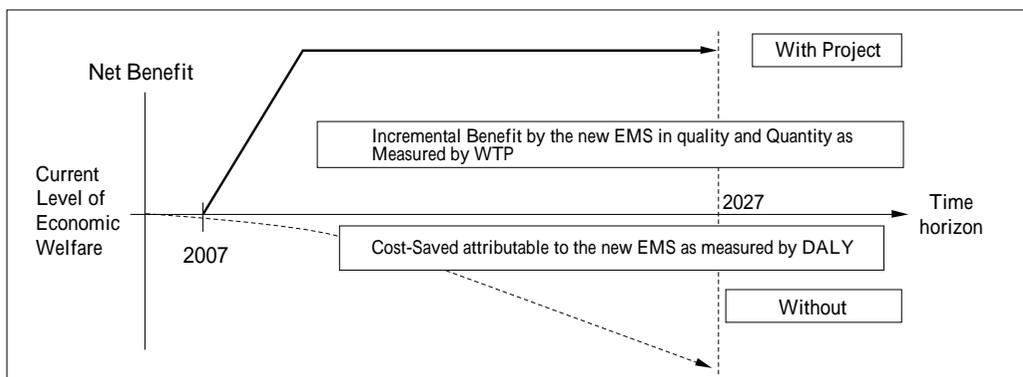
3.1.2 “ With ” and “ Without ” - Project principle

The economic benefits in the framework of “ with ” and “ without ” the project as elucidated in the above is depicted and given in Fig.2.

3.1.3 Transfer payments

Transfer payment means a shift of claims on real resources from one member or sector of society to

Figure 2 Schematic diagram of “ with ” and “ without ” project



Source: The authors

another without causing any depletion of scarce resources in the society. With this in view and acquisition, interest payments, domestic taxes and duties are eradicated from financial costs in due course of the estimation of economic costs.

3.2 Long-Run Marginal Cost (LRMC) pricing

The economic value of scarce resources used as project inputs is measured by the Long-Run Marginal Cost (LRMC) pricing method. Underlying assumption of this method is that economic benefit of investment measures is maximized when the price (benefit) is identical to the marginal cost under the assumptive “perfect competitive market”. The origins of marginal cost pricing theory date back as far as the works of P. Dupuit and subsequently H. Hotelling, in the 1930's. N. Ruggles provided a comprehensive review of work in this area up to the next decade, and the theory developed, especially for the application of in the electric power sector, with contributions from the works of M. Boiteux, P. Steiner and others from the 1950s and onwards⁴. More recently, the academic interest has led to more sophisticated investment models which permit determination of marginal costs, consideration of uncertainty, developments in peak load pricing, and so forth. On the practitioner's side, a number of contributions have been made by the economists of the international lending agencies, namely, M. Munasinghe, J. Warford, Y. Albouy, and others⁵. Backed up by these and others, the rationale for setting price equal to marginal cost to consequently attain the maximum economic welfare level is briefed herewith. The rationale for setting price equal to marginal cost may be clarified in mathematical terms as follows:

$$\text{Net Benefit (NB)} = \text{Total Revenue (TR)} - \text{Total Cost (TC)}$$

That is,

$$NB(Q) = TR(Q) - TC(Q) = p(Q) \cdot Q - TC(Q)$$

Where p and Q denote the price (the equation of demand schedule) and quantity of supply (the equation of supply schedule), respectively.

The necessary first order condition for maximizing net social benefits is to set the derivative

of the net benefit function at zero, which is mathematically derived as follows:

$$\frac{d}{dQ} NB = \frac{\partial p}{\partial Q} Q + p - \frac{\partial TC}{\partial Q} = 0$$

$$\frac{\partial p}{\partial Q} \left(\frac{Q}{p} \right) p + p - \frac{\partial TC}{\partial Q} = 0$$

$$p \left[\frac{\partial p}{\partial Q} \left(\frac{Q}{p} \right) + 1 \right] - \frac{\partial TC}{\partial Q} = 0$$

Now, the price elasticity of supply denoted as ε is defined that $\varepsilon = (\partial Q / \partial p) \times (p / Q)$. Sequentially the above equation is rewritten as below.

$$\frac{1}{(\partial Q / \partial p) \times (p / Q)} p + p = \frac{\partial TC}{\partial Q}$$

$$p \left(\frac{1}{\varepsilon} + 1 \right) = \frac{\partial TC}{\partial Q}$$

Provided that $\varepsilon = 1$ under the assumption of *perfectly competitive market*,

$$p = \frac{\partial TC}{\partial Q} = \text{Marginal Cost}$$

It is one of the basic axioms of economics that at the price p and supply (demand) Q , the total net benefit of consumption attributed to society is maximized with the optimum market clearing point (p, Q).

3.3 Economic costing and SCF

As for the foreign cost portion of the project, this procedure is reflected by way of valuating goods and services in, *notably*, CIF (Cost of Insurance and Fleet) and FOB (Fare on Board) prices, respectively. Likewise, the local cost portion of the project, a standard conversion factor (SCF) usually is applied to convert market value of local costs to its value in shadow prices. While SCF requires information on the ratios of border prices to market prices for a variety of commodities, it can be approximated by the use of data on foreign trade and net border taxes of general commodities. The approximation is provided by the border value formula as follows.

$$SCF = (M+X) / \{(M+t_m) \cdot (X+t_x)\}$$

where M and X denote the value of imports and exports in border prices, respectively, where t_m is

import duties net of subsidies and t_x is export duties net of subsidies.

In the light of EIRR analysis, the economic cost stream as per constant 2006 price level includes (i) capital cost of public intervention, and (ii) associated operation and maintenance (OM) costs for the economic life (project life) of the EMS section at CEH at a required level over the period of 2007 through 2027. To note that the estimation excluded the costs incurred prior to the afore-mentioned years as *sunk costs*.

3.4 Economic benefits

The economic benefits are due to (i) incremental EMS as measured by willingness to pay (WTP) for healthier life, and (ii) social cost-saved (recovery of foregone benefit) as measured by disability-adjusted life years (DALY).

3.4.1 Incremental benefit-willingness to pay

3.4.1.1 Sampling methodology

WTP is coherently defined as people's perceptive "*bid prices*" at the highest level for the service to be rendered, or "*proxy measurement*" of economic value. Beneficiaries (research population) in 10 cities and communities on the Black Seashore area is set at around 421,000 (58.9% of the total) in Constanta County. In carrying out the research, the double-bound dichotomous choice method of the Contingent Valuation Method (CVM) approach was applied to the interview survey.

449 residents and visitors were randomly-selected from the three categories of (i) households, (ii) business entities (commercial, industrial, and services), and (iii) beach resort visitors (expatriates and domestic travelers/residents), with the random selection procedure⁶. It would be stressed that the sampling methodology aimed at taking into account both of the people perceiving Use-Value (needs EMS) and Non Use-Value (not necessary in need of EMS now, but psychologically felt safer by the existence of EMS of avail). *CVM 2002* was used as software for statistical analysis of WTP estimation, while utilizing

*Microsoft EXCEL*⁷ to undertake a double-folded analysis.

3.4.1.2 Design of interview survey and questionnaire

Beneficiary area due to the Project includes four (4) cities and six (6) communities; the cities of Constanta, Mangalia, Navodari, and Eforie (North and South), and the communities of Corbu, Agigea, Tuzla, Costinesti, 23 August, and Limanu. Of this, the shares of interviewees amongst households, beach resort visitors, and business entities counted for 55 percent, 30 percent, and 15 percent in that order, while taking the number of the units in the Constanta County in view. With this, the actual sampling size turned out to be 250, 140, and 70 (all figures somewhat rounded) in the same order. Subsequently, all these sample figures are allocated to each of the four cities and six communities in proportion to the number of population therein. Population distribution and the sample size for each of the target cities and communities are summarized and depicted in the following Table 1.

Meanwhile, bidding games in the current research used the five-version-cascading structure, with the middle bidding price of US\$17.1 per household per month (8.0 percent of household disposable income in Romania in 2005. Neighboring bidding rates are from the lowest share of disposable income 3.0 percent (US\$ 6.4), 5.0 percent (US\$ 10.7), 8.0 percent (US\$ 17.1), 10.0 percent (US\$ 21.4), and 13.0 percent (US\$ 32.1), in the ascending order. In due course of questionnaire designing, trial-basis interview survey with the Romanian counterpart team and Focus-group meeting with around 20 randomly-sampled participants tested the validity of the questionnaire, while modifying the questionnaire as appropriate.

3.4.2 Social cost saved-DALY

Initiated by the World Bank that sponsored the aspiring study on global burden of disease (GBD) in collaboration with the World Health Organization (WHO) and the Harvard School of Public Health, the

Table 1 Number of samples by Administrative Unit and Sample Cohort

		Population	Share	Households	Beach	Entities
Share				55%	30%	15%
Sample				250	140	70
Cities	Constanta	306,860	73%	182	103	51
	Mangalia	40,805	10%	24	15	7
	Navodari	34,224	8%	20	12	6
	Eforie	9,525	2%	6	4	2
Comuna	Corbu	5,480	1%	3		1
	Agigea	5,800	1%	3	3	1
	Tuzla	6,322	1%	4		1
	Costinesti	2,396	1%	1	2	0
	23-Aug	5,201	1%	3		1
	Limanu	4,856	1%	3	2	1
Total		421,469	100%	250	140	70

Source: The authors

Disability Adjusted Life Years (DALY) has since 1993 been in a common place as an integrated measurement of population health (IMPH) for quantifying health burden. In the light of this, and as part of the economic benefits incurred (cost-saved portion) to the prospective intervention measure, economic burden of diseases and injuries in Constanta county is initially estimated by DALY and subsequently converted into monetary terms.

:

Central to the World Bank model in the estimation of DALY is the formula hereunder, with a set of parameters coherently applied to the four cases of quantity and quality of life. Some of the parameters and the constant are presumably set out including (i) age-weight modulation factor (K), (ii) constant (C), (iii) discount rate of life (r), (iv) age-weight (), and (v) disability weights (D), with numerical expression as:

$$DALY = \frac{KDCe^{-\beta a}}{(\beta + \gamma)} [e^{-(\beta + \gamma)L} (1 + (\beta + \gamma)(L + a)) - (1 + (\beta + \gamma)a)] + \frac{(1 - \kappa)}{\gamma} (1 - e^{-rL})$$

In the current study, the variable K is presumably set at one (1), as such the formula above is modified to⁸:

$$DALY = \frac{KDCe^{-\beta a}}{(\beta + \gamma)} [e^{-(\beta + \gamma)L} (1 + (\beta + \gamma)(L + a)) - (1 + (\beta + \gamma)a)]$$

In the light of the foregoing, prepositions in the current research are as follows.

- (i) Standard life expectancy at birth by gender: Male at 80.0 of age and female at 82.5 of age, while assuming a standard life table for all populations;
- (ii) Age-weighting: K is assumed to be one (1) in the current research in line with the researches previously undertaken in Romania (WB 1998), New Zealand (Ministry of Health 2001) and Australia (Ministry of Health 1999), as well as other WB documents on DALY;
- (iii) Constant (C): presumably set at 0.16243;
- (iv) Time discounting of healthy life (r): Three (3) percent of discount rate is assumed in line with the Global Burden of Disease (GBD) study⁹ and others;
- (v) Parameter for age weighting function (): Presumably set at 0.04;
- (vi) Disability weights (D): Disability adjustment factors used in the World Bank model, with the weighed average of 0.67; and
- (vii) e: an exponential.

With the above in view, the World Bank model of a “ ratio method ” of total DALY to death DALY was used with a set of disability adjustment factors based on the 2005 county mortality data collected from the National Statistic Office database and compiled by age group, gender, and residency (in aggregate 7,286 deceased)¹⁰. The human capital approach that uses loss of earnings (average household income) was applied to value social cost (benefit foregone). This provides the first order, lower-bound (conservative) estimates of foregone benefits comparing to the alternative approach to apply national income.

3.4.3 Allocative efficiency value-LRMC pricing

LRMC pricing, by mathematical definition, ensures the most efficient allocation of scarce resources, thus making the whole economy better-off. In estimation of *LRMC*-based price, the most commonly used variant of the theoretical concept in welfare economics and investment-decision theory is a levelized annuity cost coupled with recurrent cost incurred every year. In a bid to further estimate the annualized cost of investment, capital recovery factor (*CRF*), a function of (social) discount rate of capital (denoted by *i*) and economic life (*n*), will be applied in most of the places. Numerical expression of *LRMC* pricing comes in the following.

Annuitized economic cost of capital investment (Marginal Cost)

= TC (total cost) * *CRF* (*i*, *n*) + annual economic OM cost, where TC denotes the total capital investment cost,

while *CRF* is depicted as: $CRF = \frac{i(1+i)^n}{(1+i)^n - 1}$

3.5 Model configuration for analysis

Subject to technical and data/information of avail, model configured and numerical assumptions are set forth with a number of variables and parameters specified are given in the followings.

(a) Project life: 18 years with the year 2007 to commence the 3-year construction works up to 2009 and the subsequent 15-year-service period to

the year 2024;

- (b) Foreign exchange quotation: Mid-2006 market value of the RON 2.8 per US dollar;
- (c) Conversion factors: To avoid built-in market failures (non-competitive pricing, externality of the economy, etc.), the standard conversion factor (SCF) and that for unskilled labor are set at respective of 0.90 and 0.85;
- (d) Physical contingency and price contingency: physical contingency allowances are set at 5.0 percent of the base cost while reflecting unexpected increases in cost by technical uncertainty. On account of inflationary pressure, price contingencies are assumed to be 2.0 percent and 5.0 percent for the foreign and local cost components, respectively;
- (e) Financial costs and indicative investment schedule: The total financial cost is envisaged at US\$ 47.5 million, with the foreign and local cost portions respective of US\$ 18.3 million and US\$ 29.2 million. Of this, 30 percent, 40 percent, and 30 percent will consecutively be disbursed;
- (f) Operation and maintenance (O/M) costs: annual O/M cost is assumed to be 3.0 percent of the capital investment;
- (g) Salvage value: No salvage value of the fixed assets due to the project in concern is assumed in the estimation of EIRR;
- (h) Of the total 118,885 DALY estimated for Constanta 2005, share of CEH is 58.7 percent;
- (i) Of the CEH in- and out-patients, emergency medical service (EMS) beneficiaries accounts for 33.5 percent;
- (j) Of the EMS beneficiaries at CEH, 24.9 percent is the age cohort (age 15-59) for production activities in the society.

4. Results

4.1 Aggregate financial and economic costs, and economic benefits

The total financial cost of capital investment and

following O/M have been worked out to US\$ 47.5 million and US\$ 0.9 million per annum, respectively. The aggregate economic cost, likewise, turned out to respective of US\$ 31.5 million and US\$ 0.8 million as par mid-2006 price, while excluding transfer payment, taxes and duties, and price contingency from the financial cost. The economic benefits incurred to the proposed EMS in the county turned out to be US\$ 213.8 million in aggregate, with the constant economic return of US\$ 14.3 million per annum. Subsequently, net benefit of the component stands at US\$ 168.2 million in total, with the annual benefit of US\$ 13.3 million after the fourth year (2010).

4.2 Economic benefits

4.2.1 Incremental benefit: willingness to pay

4.2.1.1 Distribution of CVM-DBDC responses

YES-NO distribution for each of the five (5) bidding versions was figured out in such a way that almost one-fifth of the 449 interviewees responded in favor of “ a monthly health contribution ” regardless of the first and second bidding prices (YES-YES, 22.0 percent). Following this, YES-NO, NO-NO, and NO-YES came in the descending order while each of these accounting for 30.5 percent, 32.1 percent, and 13.6 percent, in that order. This distribution of responses by first bidding with a large proportion of YES-YES (very affirmative to payments) and a very small portion of NO-YES (a bit moderate to payments) would have reflected somewhat people's disguised

“ show-off ”, in a sense trying to “ please interviewers ”, while answering in favor of monetary devotion to healthy life. Summary of distribution of responses by bidding version is in Table 2.

4.2.1.2 WTP estimated

Willingness to Pay (WTP) as a stated preference for the “ use- and non use-“ values coherently associated with healthy-life turned out to be US\$ 17.9 per month per household (median value of 50 percent acceptance schedule), while accounting for 6.3 percent and 8.9 percent of gross and disposable incomes in Constanta, respectively. In the context of Romania as a whole (US\$ 349.3 and US\$ 248.5 in respective of gross and disposable monthly earnings), this estimate is intuitively perceived a bit high, due partly because that monthly income in the private sector is allegedly 30 to 40 percent higher than the figure in the official statistical documents.

With the monthly WTP estimate of US\$17.9 per household and the beneficiary population of 419,525 in the region, the economic benefit incurred to the incremental EMS is estimated at US\$1.6 million per annum, while the standard conversion factor (SCF) of 0.9 being applied to the financial benefit of US\$1.8 million per annum. Note the average number of household members is 2.8 persons in Romania (2005).

Table 2 Distribution of Responses (samples 449)

Version (%)	First	Second	Second	Yes-	Yes-	No-	No-	Don't	Total
Household Income	Amount (US\$)	Amount (YES)	Amount (NO)	Yes	No	Yes	No	know	
1 (3.0%)	6.4	10.7	3.2	42	37	5	12	1	97
2 (5.0%)	10.7	17.1	6.4	21	36	12	23	0	92
3 (8.0%)	17.1	21.4	10.7	13	30	12	33	6	94
4 (10.0%)	21.4	32.1	17.1	14	21	14	36	1	86
5 (15.0%)	32.1	48.2	21.4	9	13	18	40	0	80
Total				99	137	61	144	8	449
Share (%)				22.0	30.5	13.6	32.1	1.8	100

Source: The authors

4.2.2 Social cost saved: DALY by Human-Capital approach

4.2.2.1 Key findings

- (1) Disability adjusted life years (DALY) lost in Constanta is estimated at 118,825 person-years in 2005, of which premature mortality (YLL) is 71,188 while accounting for 59.9 percent.
- (2) Disaggregating by gender, males lost 73,151 DALYs (61.6%), exceeding by 60.1 percent of females DALYs of 45,675.
- (3) The leading specific cause of total disease burden was Ischaemic heart disease and stroke (8.4%), followed by, among others, hypertensive diseases (8.3%), Cerebrovascular diseases (6.9%), Mental and behavioral disorders due to alcohol (5.7%), Alzheimer disease (5.5%), Acute respiratory infectious diseases (4.5%), Chrrrhosis of liver and chronic hepatitis (3.9%), Malignant neoplasm of bronchus and lung (3.8%), Injuries to head, neck, arm and leg (3.0%), Atherosclerosis (2.1%), and others. With this, the ten leading causes of disease burden account for 52.1 percent.

The leading 10 causes of DALY in Constanta in 2005 is given below.

4.2.2.2 Social cost saved measured by DALY

The economic benefit of the recovery of benefit-foregone (loss of earnings), as borne out by DALY, is estimated at US\$ 12.7 million per annum by the formula as under:

$$\text{Economic benefit} = \text{DALYs (118,885)} \times \text{average monthly income in Constanta (US\$201.7)} \times 12 \times \text{share of CEH beneficiary (58.7\%)} \times \text{share of EMS beneficiaries (33.5\%)} \times \text{share of productive age group (24.9\%)} \times \text{standard conversion factor (0.9)}$$

4.2.3 LRMC pricing

With the discount rate of 8.0 percent over the 18 years of expected project life (with three years project implementation and operation of 15 years that follow), *CRF* being assumed to annualize the capital investment costs was estimated at 0.11, thus leading the annualized marginal cost of US\$ 3.5 million in economic terms. With this, associated with the annual recurrent (OM) cost of US\$ 0.9 million, the economic value of the project as reflected by LRMC pricing method in gross term is estimated at US\$ 4.4 million equivalent per annum and US\$ 65.9 million in total.

Table 3 Leading 10 causes of DALY in Constanta 2005

No.	Cause	DALY (years)
1	Ischaemic heart diseases of which: Acute myocardial infarction	9,983 5,300
2	Hypertensive diseases of which: Essential hypertension	9,969 9,856
3	Cerebrovascular diseases	8,218
4	Mental and behavioural disorders due to use of alcohol	6,798
5	Alzheimer disease	6,575
6	Acute respiratory infections	5,364
7	Cirrhosis of liver and chronic hepatitis	4,940
8	Malignant neoplasm of bronchus and lung	4,628
9	Injuries to the head, neck, arm, leg	3,518
10	Atherosclerosis	2,464
DALY by all conditions		118,885

Source: The National Statistic Office, 2006

4.3 EIRR and sensitivity analysis

4.3.1 EIRR analysis WTP/DALY and LRMC method

The economic feasibility as borne out by the economic rate of return (EIRR) for the case of WTP/DALY is estimated at 32.2 percent, whereas the economic net present value (ENPV) being turned out to be US\$ 102.4 million worth at the discount rate of 8 percent. Summarized is a net cash-flow of the Project, given below as Table 4 and Fig. 3.

Likewise, EIRR incurred to the alternative LRMC pricing method is elicited at 8.5 percent by LRMC pricing, with ENPV of US\$ 1.1 million worth at the social discount rate (SDR) of 8 percent.

4.3.2 Sensitivity analysis

Sensitivity analysis indicating the resiliency against project risks is undertaken with variation in

relevant parameters of (i) lower benefit by 10 percent, (ii) capital cost overrun by 10 percent, and (iii) one year delay in project completion. The results are as follows.

4.3.3 Conclusive remarks and policy evaluation

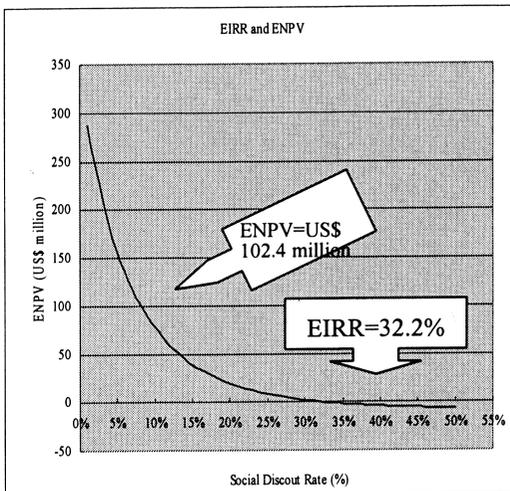
The measurement indices of EIRR and ENPV attributable to the proposed public intervention to EMS reveals economic viability high enough at 32.2 percent and US\$ 102.4 million worth (social discount rate at 8 percent), while well exceeding the generally acceptable cut-off rate of EIRR at 6-8 percent for the health sector projects. With this, the proposed project profoundly reveals supremacy being deserved for the implementation in terms of the efficient allocation of scare resources in the Romanian economy.

Table 4 Summary Net cash-flow Table (WTP/DALY)

	Investmen	OM Cost	Total Cost	WTP DALY			Total Benefit	Net Benefit
2007	9.4		9.4					-9.4
2008	12.6		12.6					-12.6
2009	9.4		9.4					-9.4
2010		0.9	0.9	1.6	12.7	14.3	13.3	13.3
2011		0.9	0.9	1.6	12.7	14.3	13.3	13.3
2012		0.9	0.9	1.6	12.7	14.3	13.3	13.3
2013		0.9	0.9	1.6	12.7	14.3	13.3	13.3
2014		0.9	0.9	1.6	12.7	14.3	13.3	13.3
2015		0.9	0.9	1.6	12.7	14.3	13.3	13.3
2016		0.9	0.9	1.6	12.7	14.3	13.3	13.3
2017		0.9	0.9	1.6	12.7	14.3	13.3	13.3
2018		0.9	0.9	1.6	12.7	14.3	13.3	13.3
2019		0.9	0.9	1.6	12.7	14.3	13.3	13.3
2020		0.9	0.9	1.6	12.7	14.3	13.3	13.3
2021		0.9	0.9	1.6	12.7	14.3	13.3	13.3
2022		0.9	0.9	1.6	12.7	14.3	13.3	13.3
2023		0.9	0.9	1.6	12.7	14.3	13.3	13.3
2024		0.9	0.9	1.6	12.7	14.3	13.3	13.3
Total	31.5	14.2	45.6	23.9	190.0	213.9	168.3	

Source: The authors

Figure 3 Economic Net-Cash-flow (WTP/DALY)



Source: The authors

Table 5 Results of Sensitivity Analysis

EIRR	Base Case	Benefit -10%	Cost+10%	1-year delay
WTP/DALY	32.2 %	29.6 %	30.0 %	29.1 %
LRMC	8.5 %	5.1 %	6.4 %	5.1 %

Source: The authors

5. Discussions (Remaining Issues)

5.1 Disability weights

Disability weights are formalized and explicitly revealed social preferences in each of the societies or cultures for different states of health, thus making likely to be different set of weights in different country. The set of disability weights applied in the current research was that for the Eastern European countries was used by the World Bank and the National Institute for Health Research and Development in their joint 1998 study, and with this in view, the evolution of disability weighting by non-fatal impairment and illness would take place to comply with general perception and culture of the Romanian people.

5.2 Age weight modulation factor - K

The World Bank model originally gives age weights (denoted as K) that vary subject to the age of onset and/or demise, while giving more than one to ages between 9-54 and below one for those in the age group of 1-8 and 55 upwards, with the maximum weight at age 25. In the current research, and almost others of close relevance that likewise estimated DALY presumably set K at one (1.0) due to an extensive efforts and ambiguity as well to assign K to each of incidence cases in the context of Romanian local city in the specific period of 2005-2006. This theoretically led the DALY estimates closer to QALYs which does not incorporate age weight modulation factor in its formula. By consequence, there would be an urgent need for standardized methodology and the specific weights to be assigned in the in the temporary Romania.

5.3 Methodological limits

DALY estimation does require extensive epidemiological data. Due partly to this constraint, the research currently in place applied a "ratio method" with the disability factors commonly used for the Eastern European countries. With this, DALY lost for mental diseases would have been either neglected or

underestimated by this methodology because these classified diseases are categorized as "non death-driven" diseases. With this and others, it would be a pressing need, in consideration of this theoretic setback and other methodological limits that the researcher had to compromise in carrying out the task, to undertake more multidisciplinary and comprehensive study on global burden of disease and injury to estimate inverse economic impact on the economy in line with the methodology as proposed by WHO/WB and other of relevance. In consequence, this will provide health policy decision makers a mechanism through which a better matching between the country's disease burden and appropriate intervention measures on an allocative efficiency basis.

Acknowledgement

The research task was jointly undertaken by the University of Tokyo and the General Directorate of Health in Constanta County, the Ministry of Health (MOH) of Romania over the period of July 2005 through September 2006. Grateful acknowledgment is made, in particular, of academic supports and cooperation provided by Dr. Moazzam Ali, the University of Tokyo, and many other researchers in Romania in the form of hearty guidance, comments and advice on the overall structure and direction of the task. The researcher's outcomes as reflected in this paper have extensively been benefited from comments provided by and discussions held with these people, while all of the editorial errors and others remain with the principal author.

Notes

- 1 In 2005, the total in-and out-patients in Romania was 4,479,847. (Source: The Public Institute of Health Research and Development, *Regularizar 2005*)
- 2 IRR, by definition, is a discount rate that equalizes discounted net benefits (benefit-cost) over the project life, and mathematically expression as follows. $IRR = r :$

- that makes $(B - C)_t \times (1 + r)^{-t} = 0$ where $(B - C)_t$ represents net benefit in the year t ($t = 1, 2, \dots, n$)
- 3 NPV $\{ (B - C)_t / (1 + r)^{-t} \}$ where $t = 1, 2, \dots, n$, and r as social discount rate.
 - 4 Dupuit, P. (1932). De l'Utilite et de sa Mesure, *La Reforma Sociale.*, Turin, H. Hotelling. (1938). The General Welfare in Relation to Problems of Railway and Utility Rates. *Econometrica*, 6, 242-269., Ruggles, N. (1949/50). The Welfare basis of the Marginal Cost Pricing Principle. *Review of Economic Studies*, 17, 29-46., and (1949-50).
 - 5 For example, see Munasinghe, M. (1984). *Guidelines for Marginal-Cost Analysis of Power System*. WB., M. Munasinghe, M. and Warford, J. (1978). *Shadow Pricing and Power Tariff Policy.*, Albouy, Y. (1983). *Marginal Cost Analysis and Pricing of Water and Electric Power*. Inter-American Development Bank, and many others.
 - 6 Number of the sample of around 450 at maximum was calculated while using the statistical software of *EPIINFO version 6* with the parameters of (i) population size of 750,000, (ii) Non-previous study alike, (iii) statistical confidential level of five (5) percent, and (iv) analytical precision level of 95 percent.
 - 7 Hearty gratitude is due to Professor K. Kuriyama at Waseda University, Japan, for his free software of *CVM* by *EXCEL, version 2* downloaded from the site: www.kkuri-mn.waseda.ac.jp
 - 8 While there seems to have been some numerical confusion in the DALY formula (Fox-Rushby, J.A. and Hanson, K. (2001). Calculating and Presenting disability adjusted life years (DALYs) in cost-effectiveness analysis. *Health policy and Planning*. Oxford University Press, 326-331.), the formula used in this paper is cited from Homede, N. (2000). *The Disability-Adjusted Life Year (DALY) Definition, Measurement and Potential Use*. The World Bank, Human Capital Development and operation Policy Working Paper.
 - 9 Murray Christopher J. L., Lopez Alan D. (1996). (Eds), *The Global Burden of Disease*. The World Bank, The World Health Organization, Harvard University.
 - 10 In so doing, the ICD-9 coding was duly converted to ICD-10 coding. In this connection, it is gratefully acknowledged to have received thoughtful academic

advises and data processing from Dr. Marcu and Ms. Galan of the Public Health Institute in Bucharest.

References

- Ali M., Miyoshi C., Ushijima H. (2006). Emergency medical services in Islamabad, Pakistan: a public-private partnership. *Public Health*, 120, 50-57.
- Asian Development Bank (2000). *Handbook for the Economic Analysis of Health Sector Projects*.
- Asian Development Bank (2003). *Economic Analysis of Health Projects: A Case study in Cambodia*. ERD Technical Note No.6.
- European Observatory on Health Care System (2000). *Health Care System in Transition, Romania*.
- Fox-Rushby J.A, and Hanson K. (2001). Calculating and Presenting disability adjusted life years (DALYs) in cost-effectiveness analysis. *Health policy and Planning*. Oxford University Press, 326-331.
- Homede N. (2000). *The Disability-Adjusted Life Year (DALY) Definition, Measurement and Potential Use*. The World Bank, Human Capital Development and operation Policy Working Paper.
- Hyder A.A., Wali S.A., Ghaffar A., Masud T.I., Hill K. (2005). Measuring the burden of premature mortality in Pakistan: use of sentinel surveillance systems. *Public Health*, 459-465.
- Mathers C., Vos T., Stevenson C. (1999). *The Burden of Disease and Injury in Australia*. Australian Institute of Health and Welfare.
- Murray Christopher J. L., Lopez Alan D. (1996). (Eds), *The Global Burden of Disease*. The World Bank, The World Health Organization, Harvard University.
- New Zealand Ministry of Health (2001). *The Burden of Disease and Injury in New Zealand*. The research applied multi-state life table methods for the estimation of incidence from prevalence, recovery and mortality data.
- Peabody J.W. et al. (2005). The burden of disease, economic costs and clinical consequences of tuberculosis in the Philippines. *Health Policy Planning*, 20 (6), Oxford University Press.
- Sassi F., Calculating (2006). Calculating QALYs, comparing QALY and DALY calculations. *Health Policy*

- Planning*, 21(5), Oxford University Press, 402-408.
- World Bank (1998). *Romania Health Sector Support Strategy*, with Annex 4A (55-69) in particular.
- World Bank (1990). *World Development Report 1990*, 26-27, Appendix B *Study MRC Policy Brief No.1*.
- World Bank (1993). *World Development Report 1993*, 26, Appendix B.
- World Bank (2004). *Deaths and Disease Burden by Cause: Global burden of Disease Estimates for 2001 by World*.
- World Bank (2004). *Derivation of the DALY Estimates*. Health Population Advisory Services *Bank Country Groups*. Disease Control Priorities Project Working Paper No. 18.
- World Bank (2005). Improving primary health care in Romania. *Output-based Contracting for Public Services*, 2005.
- World Health Organization (2005). *ICD-10 2nd Edition*.
(2007.1.30受理)

【研究論文】

社会的環境管理能力の形成プロセスに係る因果構造の分析： 都市大気汚染対策を事例にして

村上 一真 松岡 俊二 金原 達夫

広島大学大学院国際協力研究科

mkazuma@hiroshima-u.ac.jp smatsu@hiroshima-u.ac.jp t-kimbara@hiroshima-u.ac.jp

要 約

本研究は、行政、企業、市民という社会的アクターの環境管理能力により表される社会的環境管理能力に関して、アクター間の役割および各アクターの環境管理能力に一定の代替性があるとの仮説に基づき、大気汚染対策に係る社会的環境管理能力の形成プロセス、および各アクターの環境管理能力の形成プロセスに係る因果関係を実証的に明らかにした。具体的には、横浜市、名古屋市、大阪市の1971～2000年における大気汚染対策を事例に、Stepwise Chow Testを用いて社会的環境管理能力の構造変化の時期を示し、アクター間の環境管理能力に一定の代替関係が生じていることを示した。これより、社会的環境管理能力は、能力水準に応じたアクター間の役割の変化・代替を経ながら形成されていくという能力形成プロセスを都市別に示した。そして、アクター間の代替が起こるまでの期間を対象に、行政の環境管理能力の向上が、市民および企業の環境管理能力を向上させるという社会的環境管理能力の形成プロセスに係る因果関係を、構造方程式モデルにより実証した。

キーワード

社会的環境管理能力、社会的環境管理能力の形成プロセス、
構造方程式モデル、大気汚染対策

1. 研究の背景と目的

「社会的環境管理能力」は、政府・企業・市民の3つのアクター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である(松岡他 2004)。そして、この社会的環境管理能力は、行政、企業、市民社会全体の適切な参加により、環境管理を進めるという「環境ガバナンス」(松下 2002)の稼働力として位置づく(村上・松岡 2006c)。

この社会的環境管理能力の測定に関して、村上・松岡(2006b)は、村上・松岡(2006a)で探索的因子分析により示された3つの能力要素を用い、検証的因子分析モデルおよびその分析結果を踏まえた構造方程式モデルにより、社会的環境管理能力を表1に基づき、3アクター(G:行政、F:企業、C:市民)と3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・情報・技術の提供能力)で表し、大気質改善との因果関係を検証した。

また、村上・松岡（2006c）では、村上・松岡（2006b）が対象とした北九州市と大阪市に加えて、横浜市、名古屋市、神戸市の5都市を対象に、図1のモデルのように、国レベルの社会的環境管理能力の影響を考慮し、都市の大気質改善に貢献する社会的環境管理能力の構造を確認した。この分析対象の空間の拡張および都市数の拡大による検証を通じ、観測可能なデータから大気質改善に貢献する潜在的な社会的環境管理能力の水準を推定する評価基準体系として、表1の能力評価フレームの適用可能性を示した。

さらに、村上・松岡（2006d）は、汚染排出構造が類似している横浜市、名古屋市、大阪市の3都市を対象に、村上・松岡（2006c）で確認された社会的環境管理能力と、大気質および経済成長との関係性について、「経済成長を一因に形成される社会的環境管理能力が、大気質改善に貢献する」という図2の因果構造を、構造方程式モデルにより実証した。

ただし、村上・松岡（2006d）は、社会的環境

管理能力および各アクターの環境管理能力が、どのようなプロセスや因果関係で形成されるのかという動的な検証は行っていない。

ここで、環境ガバナンスやその稼働力である社会的環境管理能力は、アクター間の役割の一定の代替性を議論の前提としている。このアクター間の役割の変化・代替は、小さな政府論に関するPPP（Public Private Partnership）や、多様な主体の発意に基づく主体間での対話と相互調整による都市計画策定（小泉・西浦 2003）などの多くの分野で、公共性の担い手である多様なアクターの役割遂行として実際にみられる。ここで、各アクターには、その役割を適切に遂行できるだけの能力が備わる必要がある。したがって、各アクターの環境管理能力からなる社会的環境管理能力は、能力水準に応じたアクター間の役割の変化・代替を経ながら形成されていくと想定される。

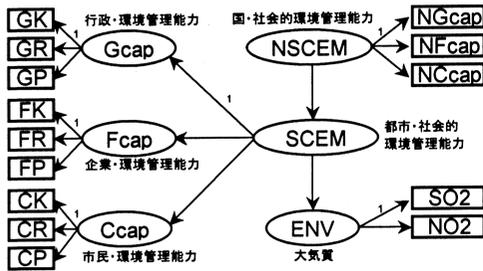
また、社会的環境管理能力は、図3のように、社会経済状態、環境質との相互関係に基づき形成され、稼働する（松岡・朽木 2003）。したがって、

表1 アクター・ファクター・マトリクス（大気汚染対策）

サイクル	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・対策課題設定(知識創造・蓄積)	
能力要素	P: 政策・対策の遂行能力 (policy & measure)	R: 環境対策資源の運用能力 (resource management)	K: 知識・情報・技術の提供能力 (knowledge & technology)	
アクター	G: 行政	<ul style="list-style-type: none"> 規制的手法 <ul style="list-style-type: none"> 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 <ul style="list-style-type: none"> 環境税、課徴金・補助金制度の整備・運用 自主的手法 <ul style="list-style-type: none"> 公害防止協定(法的根拠なし)の締結 環境学習・教育政策の推進 	<ul style="list-style-type: none"> 人員・組織 <ul style="list-style-type: none"> 環境関連部署の設置・活動 大気汚染対策組織(委員会、審議会、企業・市民協議会)の設置・活動 大気汚染対策部署職員数の拡充・活動 資金・予算 <ul style="list-style-type: none"> 大気汚染対策予算の拡充・活用 施設・設備等 大気汚染モニタリングシステムの整備・運用 大気汚染警報設備、情報システムの整備・運用 	<ul style="list-style-type: none"> 調査・研究 <ul style="list-style-type: none"> 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 <ul style="list-style-type: none"> 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	<ul style="list-style-type: none"> 規制的手法 <ul style="list-style-type: none"> 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 <ul style="list-style-type: none"> 補助金制度等の活用による対策 自主的手法 <ul style="list-style-type: none"> 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 	<ul style="list-style-type: none"> 人員・組織 <ul style="list-style-type: none"> 環境関連部署の設置・活動 環境関連部署職員数の拡充・活動 環境管理者数、公害防止管理者数の拡充・活動 資金・予算 <ul style="list-style-type: none"> 大気汚染対策予算の拡充・活用 施設・設備等 自社モニタリングシステムの整備・運用 警報設備の設置、情報システムの整備・運用 	<ul style="list-style-type: none"> 調査・研究 <ul style="list-style-type: none"> 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 <ul style="list-style-type: none"> 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	<ul style="list-style-type: none"> 規制的手法 <ul style="list-style-type: none"> 大気汚染対策に係る法規制の遵守(野焼き等) 行政の汚染監視体制の補完(SO₂の簡易測定等) 経済的手法 <ul style="list-style-type: none"> 補助金制度等の活用による対策(NGO、NPO) 自主的手法 <ul style="list-style-type: none"> 苦情、要望、ロビイング、公害防止協定への参加 エコドライブ、公共交通利用等、省エネ・省資源生活への転換 グリーン購入、エコファン 	<ul style="list-style-type: none"> 人員・組織 <ul style="list-style-type: none"> 環境対策人員数の拡充・活動(NGO、NPO) NGO、NPO活動への参加・活動 環境イベント、環境講座等への参加・活動 資金・予算 <ul style="list-style-type: none"> 環境対策関連予算の拡充・活用 施設・設備等 環境対策に係る施設・設備の確保・運用(NGO、NPO) 環境配慮製品の導入・活用(省エネ、新エネ機器) 	<ul style="list-style-type: none"> 調査・研究 <ul style="list-style-type: none"> 調査・研究(NGO、NPO) 監視・モニタリング(NGO、NPO) 情報公開・共有 <ul style="list-style-type: none"> 大気汚染状況の把握 環境関連情報の獲得 環境学習、環境教育の実施

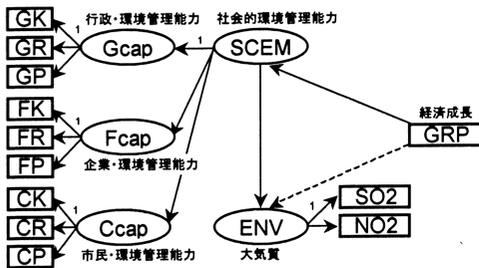
(出所) 村上・松岡（2006c）

図1 SCEM、大気質の因果関係



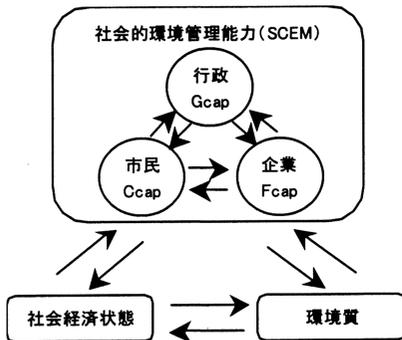
(注) SCEM : Social Capacity for Environmental Management (社会的環境管理能力)
(出所) 村上・松岡 (2006c)

図2 SCEM、大気質、経済成長の因果関係



(出所) 村上・松岡 (2006d)

図3 社会的環境管理能力の位置づけ



(出所) 松岡・朽木 (2003) より一部修正

社会的環境管理能力および各アクターの環境管理能力は、都市の社会経済状態などに一定規定されるため、その能力は社会経済状態の変化に応じて変化する。つまり、各アクターの環境管理能力は、アクター間の相互関係に基づく自己学習プロセスと、社会経済状態などの外部環境からの影響に基

づいて形成され、稼動する。そして、社会的環境管理能力の向上こそが、環境政策や国際環境協力の効率的・効果的实施に繋がる(松岡他 2004)。

本研究は、日本の都市の大気汚染対策を事例に、アクター間の役割および各アクターの環境管理能力に一定の代替性があるとの仮説に基づき、大気汚染対策に係る社会的環境管理能力の形成プロセス、および各アクターの環境管理能力の形成プロセスに係る因果関係の解明を目的とする。そして、どのアクターの環境管理能力を、どのような順序、タイミングで向上させるのが、大気汚染対策の効率的・効果的实施に繋がる社会的環境管理能力の形成となるかを示す。

本研究の構成は、まず、2で分析対象とデータ、分析手法を示す。そして、3で分析結果を示し、4で考察、5で結論を述べる。

2. 分析対象および方法

(1) 分析対象の設定

分析対象の大気汚染物質、期間、都市は、村上・松岡 (2006d) と同様とする。大気汚染物質は、工場・事業所等の固定排出源が中心の産業公害型大気汚染である二酸化硫黄 (SO₂) と、固定排出源に加え自動車等の移動排出源の寄与も大きい都市・生活型大気汚染である二酸化窒素 (NO₂) とし、排出源および解決困難度の異なる2つとする。

分析期間は、1971年から2000年までの30年間とする。これは、原嶋・森田 (1995) による環境政策の前進期 (1965~1974年)、松岡・朽木 (2003) による社会的環境管理能力の本格的稼動期 (1970~1980年代半ば) という、大気汚染対策が本格的に実施され、成果をあげ始めた時期である1970年代からを対象とすることとなる。

対象都市は、1971年時点で環境政策に一定の権限を有する政令指定都市であり、硫酸化物に係る総量規制指定地域 (1974年11月第一次指定)、および、二酸化窒素の1時間値の1日平均値が0.06ppmを超える地域 (昭和54年8月7日付 環大企310号) である5都市 (横浜市、名古屋市、大阪

市、神戸市、北九州市)のうち、村上・松岡(2006d)で示したように、汚染排出構造が類似している横浜市、名古屋市、大阪市の3都市とする。これら3都市は、多数の固定排出源と移動排出源の存在を要因に、SO₂、NO₂濃度が高く推移している(村上・松岡 2006d)。

本研究は、産業構造や都市構造に起因する都市ごとの大気汚染物質の排出構造の違いを考慮し、工業化およびモータリゼーションの進展した3都市を対象に分析し、そこでの共通的な結論を導き出す。

(2) データ

村上・松岡(2006c、2006d)と同様のデータを用いる(表2)。各アクター(G:行政、F:企業、C:市民)の環境管理能力は、それぞれ3つの能力要素(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・情報・技術の提供能力)で表すことができる。また、3つの能力要素は、表1の「サイクル」の項に示されているように、政策サイクル、知識サイクルの3区分に対応している(村上・松岡 2006a)。これより、各アクターの基本的な社会的役割、行政:政策、企業:生産、市民:消費のプロセスにおける大気汚

染対策に関する一連の事象を、3つに区分しデータ選択を行う。つまり、各アクターの環境管理能力は、SO₂およびNO₂の濃度改善というアウトカムとの関係において、その達成にむけたインプット(K、R)、アウトプット(P)にあたる3つの能力要素で表現される。

データは、表1の3×3のマトリクスに基づき、3都市ともに1971~2000年において継続的に収集可能な統計データから選択される。なお、村上・松岡(2006b)にも示したが、途上国の都市などでの実際の検討に際しては、これら統計データに加え、インタビューやアンケートによる質的情報に基づいて能力を評価する。ここでは、3都市・30年間共通に収集できる統計データにより、能力を代理的に表すこととなる。なお、各アクターのデータは、行政は市および環境科学研究所あたり、企業は事業所あたり、市民は世帯あたりに加工し、アクターごとに単位を統一することで指標化する。

行政の環境管理能力(Gcap)

行政の環境管理能力は3つの指標で表される。まず、大気汚染対策に関する「政策」のアウトプットとして、「大気汚染政策の適切な遂行(gp)」が設定できる。そして、この達成の背景には、

表2 データ

項目	データ	出所
Gcap (行政)	gk GK	政策立案・実施の基礎となる科学的研究の実施 :市立の環境科学研究所の研究成果に係る論文の掲載数
	gr GR	環境科学研究所報(各都市、各年版) 環境白書等(各都市、各年版) 厚生省「衛生行政業務報告」(各年版) 地方公務員給与と制度研究会「地方公務員給与の実態」(各年版)
	gp GP	大気汚染政策の遂行 :市および環境科学研究所による工場・事業場等への大気汚染検査件数
Fcap (企業)	fk FK	設備に体化された資源効率的な生産に係る技術・ノウハウの蓄積 :製造業部門の有形固定資産の年末現在高(実質値)
	fr FR	経済産業省「工業統計表」(各年版) 産業環境管理協会「産業公害」,「環境管理」(各年版) 内閣府「国民経済計算 需要項目別時系列表 年度デフレータ」
	fp FP	資源効率的な生産 :製造業部門の製造品出荷額/原材料使用額
Ccap (市民)	ck CK	環境意識醸成などに係る知識・情報化 :通信および書籍・他の印刷物に係る支出額(実質値)
	cr CR	総務省「家計調査年報」(各年版) 地方財政調査研究会「地方財政統計年報」(各年版) 総務省「平成12年基準 消費者物価連続指数」 経済企画庁「国民経済計算年報」(各年版)
	cp CP	個人の環境意識を環境行動に促す「場」の創造 :市の社会教育費(実質値)
ENV (大気質)	SO ₂	公共交通機関の利用 :バス・電車の通学および通勤定期に係る支出額(実質値)
	NO ₂	二酸化硫黄(SO ₂)濃度(ppm)(一般測定局における年平均値) 二酸化窒素(NO ₂)濃度(ppm)(同上)
GRP (経済成長)	1人当たり市内総生産(実質値)	環境省「日本の大気汚染状況」(各年版) 内閣府「県民経済計算年報」(各年版)

(出所)筆者作成

「政策立案・実施の基礎となる科学研究 (gk)」、
「科学研究を踏まえた政策化を推進するための
組織力 (gr)」が存在し、gk gr gpという大気
汚染に対する一連の政策プロセスが導かれる。こ
こで、これらに対応する能力要素を、村上・松岡
(2006c)と同様に、表2にあるGcapのGK、GR²、
GPの3つのデータで代理させる。

松下(2002)は、環境ガバナンスに関連して、
政策プロセスを、課題設定、政策形成、政策実施
に区分し、環境問題の存在と行政的課題としての
認識、政策決定と現実の政策実施などにギャップ
があり、その要因把握のために政策プロセス区分
ごとの検討の必要性を述べている。したがって、
科学研究、政策化に係る組織体制、政策遂行を、
政策プロセス区分に対応する個別の能力要素とし
て設定することは妥当といえる。

企業の環境管理能力 (Fcap)

企業の環境管理能力は3つの指標で表される。
まず、大気汚染対策に関する「生産」のアウトプ
ットとして、「資源効率的な生産 (fp)」が設定で
きる。そして、この達成の背景には、クリーナー
・プロダクション技術等による「資源効率的な
生産に係る技術・ノウハウの蓄積 (fk)」、
「公害防止や生産工程管理のための組織体制 (fr)」
が存在し、fk fr fpという大気汚染に対する一連
の生産プロセスが導かれる。ここで、これらに対
応する能力要素を、村上・松岡(2006c)と同様
に、表2にあるFcapのFK、FR、FPの3つのデー
タで代理させる。なお、FK「有形固定資産の年末
現在高」は、民間企業設備デフレータにより実質
化する。

ここで、FKは生産活動全般に係る資産であり、
クリーナー・プロダクション技術のみを範囲とす
るわけではないが、企業は資源効率的な生産に係
る設備の導入とその運用に係る技術・ノウハウの
蓄積という合理的な行動をとると想定できるため、
これに代替する。

また、FRは、全国9会場別に示されており、
「(会場別合格者数) × (当該都市製造業事業所
数 / 当該地方経済産業局の管轄地域内製造業事業
所数)」により推計した。そして、1971年の第1回
試験合格者の平均年齢が約30歳であり、公害防止

に係る知識・ノウハウを持った人材が企業内で継
続して貢献しているものとして累計値とする。

紺野・野中(1995)は、能力ベース経営におい
て、能力を組織資源レベル(知識資源、認知的能
力)、知識変換レベル(システム、プロセス、ス
キル)、プロダクトレベル(コアプロダクト、補
完的製品・サービス)の3層に分類し、それらが
相互に関連することで、企業の総体的な能力が実
体化していると述べており、本研究の企業の能力
要素K、R、Pの分類と整合する。

市民の環境管理能力 (Ccap)

市民の環境管理能力は3つの指標で表される。
まず、大気汚染対策に関する「消費」のアウトプ
ットとして、「公共交通機関の積極的な利用 (cp)」
が設定できる。そして、この達成の背景には、
「環境意識醸成などに係る知識・情報化 (ck)」、
「個人の意識を環境行動に促す『場』の創造 (cr)」
が存在し、ck cr cpという大気汚染に対する一
連の消費プロセスが導かれる。ここで、これらに
対応する能力要素を、村上・松岡(2006c)と同
様に、表2にあるCcapのCK、CR、CPの3つのデー
タで代理させる。なお、CK、CPは都市別品目別
デフレータ、CRは教育支出デフレータで実質化
する。

ここで、CKは、環境に関する考えや価値観、
基本的能力の水準を代替する³。ここで、廣松
(1986)の情報化指標作成に係る情報流通構造の
分類であるパーソナル/マス・コミュニケーション
区分を踏まえる。電話、はがき等の通信費をパ
ーソナル・コミュニケーション、新聞、雑誌等の
書籍・他の印刷物費をマス・コミュニケーション
の代表として、これらの和を指標とする。

また、CRは、公民館費、図書館費、青少年教
育施設費、女性教育施設費、社会教育活動費など
を含むものであり、公的な係りに関して、社会規
範、慣習・習慣の維持に寄与する⁴。また、教育
の外部性、つまり個人に教育需要の決定を委ねる
と社会全体にとって望ましい教育水準が達成され
ない可能性が高いため(小塩 2002)、社会教育は
個人の知識・情報化を補完し、集団での活動を促
進させる機能を有す。これらに関して、知識は特
定の時間、場所、他者との関係性の中で創発、修

正されるものであり(野中他 2003) この『場』の創造に社会教育は貢献するものといえる。つまり、個々人の家庭での学習に基づく環境意識は、他者との意見交換や情報共有により、環境行動につながる水準にまで高まる。これより、積極的な公共交通機関の利用という消費行動の変化(環境行動)が期待されることとなる。

CPIに関しては、表1にあるように、公害防止協定への参加、行政の汚染監視体制の補完なども候補にあがる。ただし、これらはデータ制約とともに、固定排出源に加えて、自動車排ガスの寄与が大きいNO₂に対する環境行動を含むものとは言えない。したがって、自動車の利用者でもある市民の環境管理能力は、NO₂だけでなくSO₂を含む大気汚染全般に対する意識の高まりとして、「公共交通機関の積極的な利用」という環境行動に結びつく捉える。例えば、昨今のノー・マイカー・デーやパーク・アンド・ライドの推進状況は、市民の環境意識の高まりによる環境行動という、環境管理能力により規定される。

Stern(2000)は、個人の環境行動とその要因に関する先行研究を整理し、環境行動は、「1. 環境活動組織への参加」、「2. 環境イベント参加や環境政策への賛同」、「3. 低環境負荷の財・サービスの購入・消費」、「4. 企業、団体等の構成員としての環境行動等」の4つの視点から捉えることができるとしている。

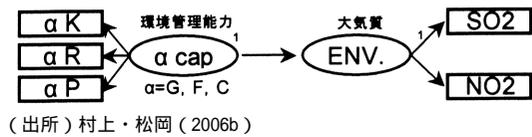
一方、これらの環境行動の要因を、「a. 道徳、信念、価値観などの考え」、「b. 教養などの基本的な能力、応用的な知識・技術」、「c. 経済性、法規制、社会規範などの外的要因」、「d. 慣習・習慣」に4分類し、複数の要因の相互作用により、特定の環境行動が導かれることを認識すべきとしている。

本研究では、環境行動は「3. 低環境負荷の財・サービスの購入・消費」としての「公共交通機関の積極的な利用(cp)」となり、その要因は、a、bにあたる「環境意識醸成に係る知識・情報化(ck)」、c、dにあたる「個人の意識を環境行動に促す『場』の創造(cr)」となる。

社会的環境管理能力(SCEM)

各アクターの環境管理能力および社会的環境管理能力の指標化を示す。村上・松岡(2006a)は、

図4 アクター別構造方程式モデル



探索的因子分析での因子負荷量の寄与率⁵を、能力要素ごとのウェイトとして捉え、3つの能力要素P、R、Kのウェイトを31.8、47.3、7.5とした。この結果をもとに、村上・松岡(2006a)と同様に、各アクターの環境管理能力は、上述した ~ の各アクターの3つ能力要素(P、R、K)を表すデータを、このウェイトにより加重平均した値として設定する。

社会的環境管理能力の指標化に係るウェイトは、村上・松岡(2006b)の分析結果を用いる。村上・松岡(2006b)は、図4のモデルに基づき、各アクターの環境管理能力の指標化を、上述した ~ の各アクターの3つ能力要素(P、R、K)を表すデータを、このウェイトにより加重平均した値として設定する。

大気質(ENV.)に係る標準化係数は、-0.86、-0.94、-0.91となった。この各アクターの環境管理能力の大気質改善への貢献度の絶対値を、アクターのウェイトとする。したがって、社会的環境管理能力は、能力要素ウェイトにより加重平均して得られた各アクターの環境管理能力を、このアクター・ウェイトにより加重平均した値として設定する。

SO₂濃度、NO₂濃度は、図5、図6のように一般測定局における年平均値とする。また、経済成長(GRP)は、1人あたり市内総生産とし、GDPデフレータで実質化する。

(3) 分析手法

まず、各アクターの環境管理能力および社会的環境管理能力の推移を踏まえ、Stepwise Chow Testおよび相関分析を用いて、アクター間の環境管理能力の代替性に係る検証を行い、大気汚染対策に係る社会的環境管理能力の形成プロセスを明らかにする。これを踏まえて、構造方程式モデルにより、社会的環境管理能力を構成する各アクターの環境管理能力の形成プロセスに係る因果関係

図5 SO₂濃度の推移

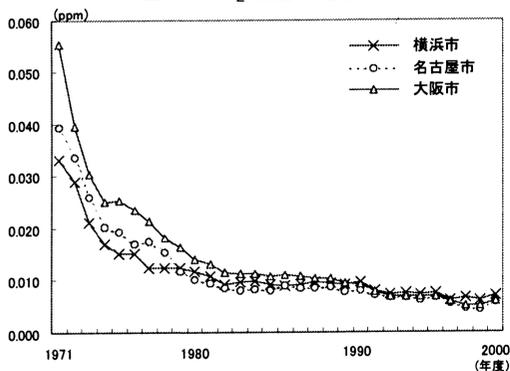
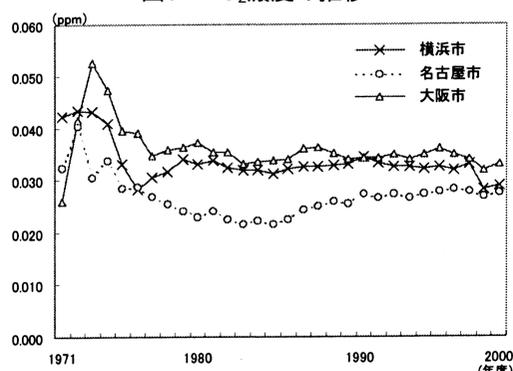


図6 NO₂濃度の推移



(出所) 環境省「日本の大気汚染状況」(各年版)、環境白書等(各都市、各年版)

を検証する。そのため、2つの実証分析を行う。

分析1：アクター間の代替性の検証

Stepwise Chow Testを用いて、行政、企業、市民の環境管理能力からなる社会的環境管理能力の構造変化の時期を明らかにする。そして、構造変化前後の期間それぞれでのアクター間の相関関係を考察することで、アクター間の環境管理能力の代替性を明らかにする。この相関関係と代替性について、アクター間の相関係数の符号に着目し、負符号であれば、アクター間の環境管理能力に代替性が生じていると解釈する。つまり、本研究での「代替」は、社会的環境管理能力における各アクターの環境管理能力の配分割合の変化が、代替関係となるかを意味する。

ここでは、社会的環境管理能力の構造変化を、それを構成する各アクターの役割の変化・代替に伴う環境管理能力の水準変化に拠るものとして捉える。つまり、部分としての各アクターの環境管理能力水準のバランスの変化により、全体としての社会的環境管理能力に構造変化が起こると捉える。これにより、社会的環境管理能力は能力水準に応じたアクター間の役割の変化・代替を経ながら形成されていく、という能力形成プロセスの全体像を示すこととなる。

なお、Chow Testは、時系列データを構造変化が起こったと想定される時点を事前に決めて、前期、後期に2分割し、両期間で推定されるパラメータが等しいかどうかをF検定により検定し、パ

ラメータが有意に異なれば構造変化が生じていると判定する方法である(高林 1988)。また、Chow Testの応用であるStepwise Chow Testは、構造変化の時期が事前に確定されていない場合に用いる方法で、分割時点を1期ずつ後方にずらしながら、それぞれでF値を算出し、構造変化期を特定する方法である(二宮 1977)。

分析2：アクター間の因果関係の検証

構造方程式モデル(Structural Equation Model: SEM)を用いる。構造方程式モデルは、直接観測できない潜在変数を導入し、その潜在変数と観測変数との間の因果関係を同定することにより、社会現象や自然現象を理解するための統計的アプローチである(狩野・三浦 2002)。

ここでは、村上・松岡(2006d)で明らかにした社会的環境管理能力、大気質、経済成長の因果構造を踏まえ、大気質、経済成長との関係性の中で、社会的環境管理能力を構成する各アクターの環境管理能力の形成プロセスに係る因果関係の検証を行う。

以上、分析1、2より、大気汚染対策に係る社会的環境管理能力の形成プロセスの検証と、その形成プロセスの中身である各アクターの環境管理能力の形成プロセスに係る因果関係の検証がなされる。

3. 分析結果

(1) 分析1：アクター間の代替性の検証

各都市の1971～2000年のアクターごとの環境管理能力の推移は図7、8、9となる。データは標準化し（平均0、標準偏差1）推移の傾向を捉えるため、2次曲線の近似曲線を示した。近似曲線は、3都市ともに、行政、市民は上に凸、企業は下に凸の形状となった。また、1971～2000年の期間において、行政では横浜市、大阪市、市民では横浜市、名古屋市で2次曲線に極大点がみられる。

各アクターの環境管理能力は、グラフおよび近似曲線の形状より、3都市ともに、まず、行政および市民の環境管理能力の向上が見られる。その後、企業の環境管理能力が向上するにしたいが、行政および市民の環境管理能力が横ばいあるいは低下傾向になることが、おおまかに見て取れる。

また、各都市の1971～2000年の社会的環境管理能力の推移は図10となる。2次曲線の近似曲線は、3都市ともに上に凸の形状となった。3都市それぞれの社会的環境管理能力は、グラフおよび近似曲線の形状より、1971～2000年の期間において極大点を迎えていない右肩上がりの推移となっている。

以上より、3アクターの環境管理能力からなる社会的環境管理能力は、それを構成する各アクターの環境管理能力が上下するにも関わらず上昇傾向にあるといえる。これより、アクター間の環境管理能力の配分割合の変化として、アクター間の環境管理能力の代替可能性が示唆される。次に、これをStepwise Chow Testを用いて、実証的に確認する。

まず、上述してきたような図7～10のグラフおよび近似曲線の形状に係る考察、および表3に示した社会的環境管理能力の発展ステージ・モデル⁶を踏まえて、Stepwise Chow Testで検証するモデルを設定する。表3の「システム形成期」では、行政は自身の能力形成とともに、市民の環境意識の醸成、企業への社会的圧力の形成の役割が指摘されている。「本格的稼働期」では企業の能力形成の重要性が示され、「自律期」では企業、市民が環境管理のイニシアティブを担う、という能力形成プロセスが記述的に示されている。また、こ

図7 各アクターの環境管理能力推移（横浜市）

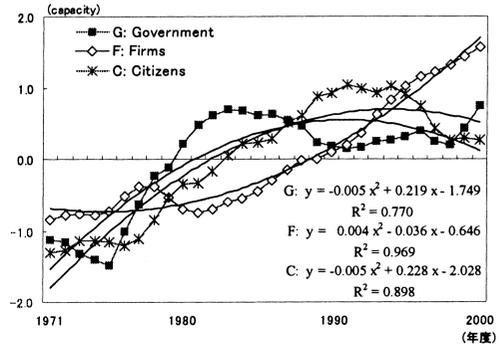


図8 各アクターの環境管理能力推移（名古屋市）

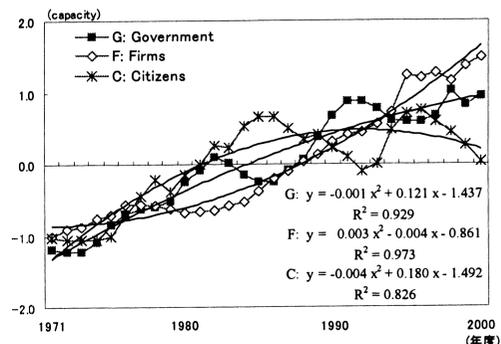


図9 各アクターの環境管理能力推移（大阪市）

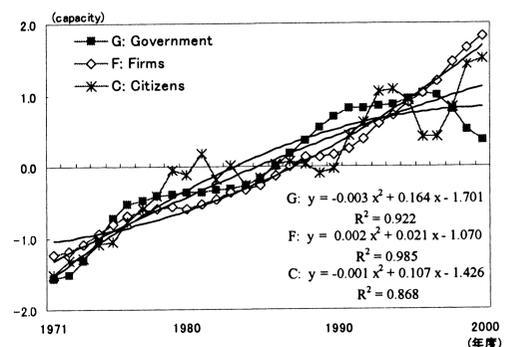
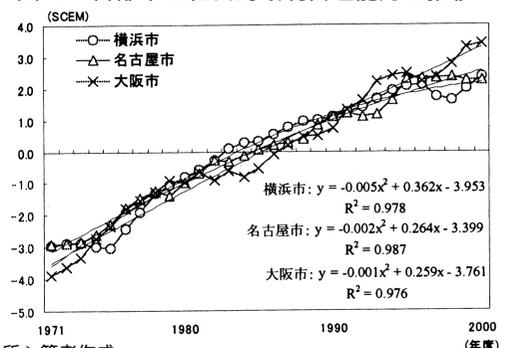


図10 各都市の社会的環境管理能力の推移



(出所) 筆者作成

表3 社会的環境管理能力の発展ステージ（大気汚染対策）

ステージ	概要
システム形成期 ↓	<ul style="list-style-type: none"> 環境管理に関する制度の束である社会的環境管理システムが形成される この時期は、「政府（行政）の能力」形成が重要 一般的には、環境法の整備（法制度整備）、環境行政制度の整備（環境省などの整備）、環境情報制度（汚染モニタリング・ネットワーク整備、データの収集と活用・公開）の整備という3つのステップがある 特に環境情報においては、環境モニタリング・ステーション数のみでなく、情報ネットワークの整備と公開による市民の環境意識の醸成と汚染者（企業）への社会的圧力の形成が重要
本格的稼働期 ↓	<ul style="list-style-type: none"> 汚染削減を本格的に行っていく時期であり、汚染が増加傾向から減少傾向に転じ、いわゆる環境クズネツ曲線の転換点が観察される この時期は、企業における汚染対策を可能とする技術・知識の開発・普及であり、企業における公害対策を可能とする人材・組織的資源の蓄積が最も重要 政府における明確な環境対策とその実施、市民の環境問題への関心・監視と政府の政策への支持も、もちろん重要
自律期	<ul style="list-style-type: none"> 政府・企業・市民という社会的アクターの相互関係が強くなり、社会的能力が自律的に発展していき、総合的な環境管理が行われていく 民間（企業、市民）が環境管理のイニシアティブを担い、新しく発生する環境問題への柔軟な対応が可能となる 環境政策の手法としては、直接規制から市場的手法や自主的手法の重視へと移行していく 途上国においては、従来のODAを主とした先進国との垂直的な協力関係から、民間交流などによる水平的な協力関係へと移行し、援助から卒業する

（出所）松岡・朽木（2003）松岡他（2004）より作成

れに関連して、今村（2002）は、標語としての「ガバメントからガバナンスへ」は、ガバナンスとしての多様なアクターの参加による連携・協働の中では、ガバメントのありようが改めて問われていると指摘し、「ガバメントもガバナンスも」として、行政のガバナンスにおける役割の重要性を述べている。

これらより、行政から企業および市民への役割の交代・代替により行政の環境管理能力の水準が変化することで、3アクターの環境管理能力からなる社会的環境管理能力の構造変化が引き起こされると捉える。つまり、Stepwise Chow Testのモデルを、 $SCEM_{it} = \alpha + \beta Gcap_{it} + U_{it}$ （ U_{it} : 誤差）という行政の環境管理能力を説明変数とする回帰式として設定し、社会的環境管理能力の構造変化の時期を明らかにする。

Stepwise Chow Testより、表4のとおり、Takeuchi（1991）、山本・Zhai（1995）と同様に、最も高いF値をもたらす時期を構造変化の時期と判定した。これは、図7～9より、3都市ともGcapとFcapの交点に近い時期となっている。

次に、表4で示された社会的環境管理能力の構造変化時期で区分される期間ごとに、アクター間の環境管理能力の相関係数を示した（表5、6、7）。社会的環境管理能力の構造変化が起こる前、つ

表4 Stepwise Chow Testの結果

都市	F値	年度
横浜市	84.262 **	1988
名古屋市	13.816 **	1995
大阪市	55.935 **	1994

（注）最も高いF値のみを示した

** 1%有意、* 5%有意

（出所）筆者作成

まり、アクター間の代替が起こる時期までは、3都市ともにアクター間の相関係数は全て正符号となり、統計的にも多くが有意となっている。また、各アクターの環境管理能力と社会的環境管理能力の相関係数も、全て統計的に有意な正符号となっており、この時期、全てのアクターの環境管理能力は上昇傾向となることが確認された。松岡・朽木（2003）によると、この時期は、表3のシステム形成期・後期あるいは本格稼働期・前期から、自律期への移行期にあたり、各アクターともに環境管理能力を向上させ対策を進めることで、図5、6にあるように、SO₂濃度、NO₂濃度を改善させてきたといえる。

一方、構造変化後では、アクター間の環境管理能力は、正符号で統計的に有意な相関関係を示しているものではなく、逆に、負符号が多く、統計的

表5 アクター間の環境管理能力の相関係数（横浜市）

	1971 - 1987年度(構造変化・前)				1988 - 2000年度(構造変化・後)			
	Gcap	Fcap	Ccap	SCEM	Gcap	Fcap	Ccap	SCEM
Gcap	1				1			
Fcap	0.472	1			0.540	1		
Ccap	0.894 **	0.436	1		-0.350	-0.648 *	1	
SCEM	0.969 **	0.595 *	0.954 **	1	0.767 **	0.837 **	-0.239	1

表6 アクター間の環境管理能力の相関係数（名古屋市）

	1971 - 1994年度(構造変化・前)				1995 - 2000年度(構造変化・後)			
	Gcap	Fcap	Ccap	SCEM	Gcap	Fcap	Ccap	SCEM
Gcap	1				1			
Fcap	0.861 **	1			0.337	1		
Ccap	0.687 **	0.531 **	1		-0.137	-0.251	1	
SCEM	0.953 **	0.885 **	0.836 **	1	0.701	0.619	0.382	1

表7 アクター間の環境管理能力の相関係数（大阪市）

	1971 - 1993年度(構造変化・前)				1994 - 2000年度(構造変化・後)			
	Gcap	Fcap	Ccap	SCEM	Gcap	Fcap	Ccap	SCEM
Gcap	1				1			
Fcap	0.975 **	1			-0.803 *	1		
Ccap	0.805 **	0.823 **	1		-0.788 *	0.312	1	
SCEM	0.971 **	0.975 **	0.919 **	1	-0.944 **	0.698	0.893 **	1

(注) ** 1%有意、* 5%有意

(出所) 筆者作成

に有意な相関係数もみられる。表7より、大阪市では、行政から企業および市民への役割の代替により、行政と企業および市民の環境管理能力に代替関係が生じていることが、負の相関係数より示された。

ここで、表3の自律期に、「政府・企業・市民という社会的アクターの相互関係が強くなる」、「民間（企業、市民）が環境管理のイニシアティブを担う」とあるように、表4の3都市の社会的環境管理能力の構造変化、つまりアクター間の代替は、自律期に移行した後に生じており、表3と整合する。

また、「総合的な環境管理」、「新しく発生する環境問題への柔軟な対応」とあるように、行政の役割は、自律期ではCO₂や有害化学物質、廃棄物などの解決困難な新たな環境問題への対策、そして、そのための能力形成に重心が移っていく。つまり、SO₂、NO₂という固定排出源を主な汚染源とする産業型公害については、環境基準の一定の

達成を踏まえて、能力水準が向上してきた企業の自主的な対策と、表1にある市民による行政の汚染監視体制の補完としてのWatch dog（監視者）などに一定任せて、行政は汚染行動の規制や監視を、選択的そして効率的に実施することが可能となる。これは小さな政府への転換という実際にも通じる。この役割の変化に伴う能力の要求水準の変化として、行政の大気汚染（SO₂、NO₂）に係る環境管理能力は、横ばいあるいは低下傾向となって表れると解釈できる。つまり、単に行政の環境管理能力が低下したのではない。

一方、企業の環境管理能力は、指標化において、資源効率的な生産、およびそれに係る技術・ノウハウの蓄積や生産工程管理などのための組織体制の整備という一連のプロセスとして捉えている。したがって、企業の大気汚染（SO₂、NO₂）に係る環境管理能力は、経済合理性に基づく資源効率性の継続的な追求としての環境対策の推進によ

り、右肩上がりに表れると解釈できる。

ここで、表5、6の横浜市、名古屋市の分析結果であるが、大阪市とは、企業の環境管理能力の上昇傾向は同じだが、行政、市民の環境管理能力の推移とそれに関する相関関係が異なる。CcapとFcapは負の相関となり、GcapとFcapは正の相関となっている。また、GcapとSCEMも正の相関となっている。図7、8より、行政の環境管理能力が、それまでは横ばいあるいは低下傾向だったのが1997、98年あたりから再向上し、市民の環境管理能力が横ばいあるいは低下傾向となっていることから、市民の環境管理能力の再代替として、行政の環境管理能力が再向上したと捉えることができる。また、企業の環境管理能力の上昇も、市民の環境管理能力を代替している。

ここで、市民の環境管理能力の指標は、環境意識の高まり（K、R）による環境行動（P）というプロセスを、「公共交通機関の積極的な利用」という環境行動に結びつくプロセスとして捉えている。これより、特に近年の市民の環境管理能力の推移の横ばいあるいは低下傾向は、自動車の快適性・利便性向上に対する嗜好の変化や生活スタイルの変化、都市の拡散と交通インフラ整備状況などの社会経済状態の変化の影響が大きいと想定される。その結果が、表8にあるように、1990年以降の自動車保有台数、平均交通量、平均混雑度に表れる。横浜市では、表5でCcap・SCEM、Ccap・Gcap、Ccap・Fcapが負の相関となり、自動車保有台数、平均交通量および平均混雑度が、1990年以降も、ともに増加する。名古屋市では、表6でCcap・SCEMは正の相関だが、Ccap・Gcap、Ccap・Fcapは負の相関となり、自動車保有台数が増加する。一方、大阪市は、表7でCcap・SCEMは統計

的に有意な正の相関となり、市民の環境管理能力は上昇傾向として表れ、1990年以降、表8の数値は全てが減少する。

以上より、横浜市、名古屋市では、生活スタイルの変化や交通インフラ整備を中心とした社会経済状態の影響を強く受け、市民の大気汚染（SO₂、NO₂）に係る環境管理能力は、横ばいあるいは低下傾向となって表れる。そのため、自動車交通動向の悪化に繋がり、その対応として、行政の環境管理能力が再向上していると解釈できる。そのため、大阪市のCcapとGcap間の負符号の意味とは異なる。各アクターの環境管理能力は、アクター間の相互関係に基づく自己学習プロセスと、社会経済状態などの外部環境からの影響に基づいて形成されることから、これらの能力形成に与える影響度の違いにより、都市ごとに異なる結果が生じるといえる。

以上の分析結果より、表4の時期を境に、都市ごとに代替状況は異なるが、各アクターの環境管理能力に一定の代替関係が生じていることを示した。

(2) 分析2：アクター間の因果関係の検証

分析1では、社会経済状態に影響を受けつつ、能力水準に応じたアクター間の役割の変化・代替を経ながら、社会的環境管理能力が形成されるという能力形成プロセスの全体像を都市ごとに示した。また、各アクターの環境管理能力の代替状況は、都市ごとに異なることも示した。

ここでは、分析1での社会的環境管理能力および各アクターの環境管理能力の形成プロセスを踏まえ、その形成プロセスにおいて、どのような因果関係で各アクターの環境管理能力が形成されて

表8 自動車保有台数、平均交通量および平均混雑度（一般道路計）の推移

	自動車保有台数(台)			DID・12時間平均交通量(台/12h)			DID・12時間平均混雑度		
	横浜市	名古屋市	大阪市	横浜市	名古屋市	大阪市	横浜市	名古屋市	大阪市
1971年	330,676	384,817	603,547	17,879	25,197	29,951	1.05	0.84	0.95
1980年	684,549	781,210	718,755	19,738	20,575	24,981	1.06	0.87	1.04
1990年	1,232,799	1,134,834	939,728	23,879	21,435	25,620	1.23	1.04	1.36
1999年	1,426,354	1,261,568	913,285	24,797	20,345	23,799	1.24	0.99	1.07

(注) DID：人口集中地区、平均交通量 = (交通量 × 区間延長) の合計 / 区間延長の合計
平均混雑度 = 交通量 / 交通容量 (走行台キロの合計 / 容量台キロの合計)

(出所) 国土交通省「道路交通センサス」(各年版)、大都市統計協議会「大都市比較統計年表」(各年版)

きたかを明らかにする。

なお、ここでは、構造変化前の期間、つまりアクター間の代替の起こる前の期間を分析対象とする。図7～10および表4～7に基づく分析1の結果からも明らかなように、構造変化期を境に変化する各アクターの環境管理能力の推移は、当然、アクター間の因果関係も変化させる。そのため、全期間を通じた分析ではなく、2期間それぞれで分析する必要がある。ここで、都市ごとの各アクターの環境管理能力の構造変化後の代替状況は異なるため、因果関係も都市ごとに異なることと、構造変化後のデータ数が限定されるという分析手法上の課題より、構造変化前、つまりアクター間の代替前の期間のみを分析対象とし、3都市に共通的な結論を導くこととする。なお、表4でのアクター間の代替時期は、表3の自律期にあたり、「援助から卒業する」時期と示されている。国際環境協力を踏まえると、援助の必要があるアクター間の代替前の期間において、各アクターの環境管理能力の形成プロセスに係る因果関係を明らかにすることは、援助対象とすべきアクターの選択や順序、タイミングなどの知見を示すことにも繋がる。

構造方程式モデルによる分析結果は図11となる。なお、社会的環境管理能力の構造変化前という分析時期の設定により、各都市のデータ数が限られるため、3都市のプールデータ(N=64)により分析を行った。また、モデルの識別性確保のため⁸⁾、ENV SO₂を1に固定する制約を課すことで解を求めた。

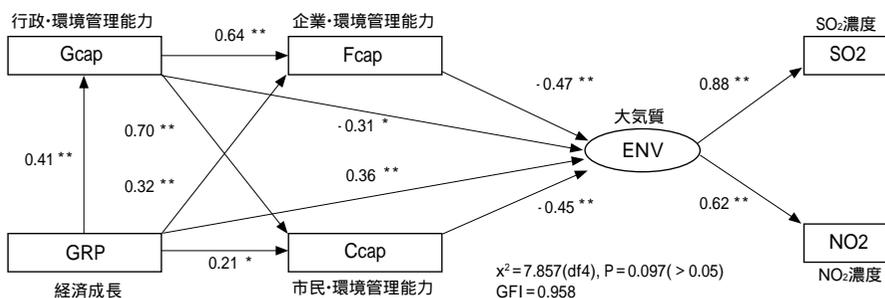
モデルの因果構造の設定は、表3のシステム形成期で、行政は自身の能力形成とともに、市民の環境意識の醸成、企業への社会的圧力形成の役割の指摘を踏まえた。また、金原・金子(2005)による、環境規制などの行政からの圧力、および環境志向などに基づく地域社会からの要請が、企業の環境経営を促進させるという因果関係モデルの提示も考慮した。なお、CcapとFcapの間のパスは統計的に有意とならなかったため省略した。また、GcapとCcapのパスの向きは、Gcap Ccapという因果関係とした。これは、上述した表3での記述とともに、仮にGcap Ccapとした場合、村上・松岡(2006d)で示されたGRP Ccapのパスが統計的に有意でなくなるという分析結果に拠る。

モデルの適合度は、 $\chi^2 = 7.857 (df4)$ 、 $P = 0.097 (>0.05)$ 、 $GFI = 0.958$ となり、一定の適合度を示した。ここでの²⁾検定は、帰無仮説を「モデルが正しい」とおくため、P値が0.05以上で仮説が棄却されないことで、5%の有意水準でのモデルの採択がなされる。また、パス係数は全て統計的に有意となった。

各アクターの環境管理能力の形成プロセスに係る因果関係は、Gcap FcapおよびGcap Ccapとして統計的な有意性が得られた。これより、行政は自身の能力形成を進める中で、市民の環境意識の醸成、企業への社会的圧力の形成を推し進め、市民、企業の能力形成を促進させたといえる。

また、各アクターの環境管理能力と大気質(ENV)の因果関係について、各アクターともに

図11 各アクターの環境管理能力の形成プロセスにおける因果関係



(注) 誤差変数、攪乱変数は省略した

** 1%有意、* 5%有意

(出所) 筆者作成

負符号で統計的に有意となり、各アクターの環境管理能力が向上することにより、SO₂濃度およびNO₂濃度が低下するという関係が示せた。

そして、経済成長（GRP）と各アクターの環境管理能力の因果関係について、各アクターともに正符号で統計的に有意となり、経済成長に伴い各アクターの環境管理能力が向上するという関係が示せた。

さらに、経済成長（GRP）と大気質（ENV）の因果関係について、GRP ENVは正符号で統計的に有意となり、経済成長に伴い大気質が悪化するという関係となった。これより、経済成長に伴い転換点を経て環境質が改善していく逆U字を描く環境クズネツ曲線仮説に関しては、村上・松岡（2006d）のように、経済成長それ自身が直接大気質改善に寄与しているのではなく、経済成長を一因に形成される社会的環境管理能力および各アクターの環境管理能力が、大気質改善に寄与していると捉えるのが適切であるといえる。

以上、社会的環境管理能力および各アクターの環境管理能力の形成プロセスに係る因果関係として、行政の環境管理能力の向上が、市民および企業の環境管理能力を向上させるという因果関係を実証的に明らかにした。

4. 考察

以上の分析結果を踏まえて、大気汚染対策の効率的・効果的实施に繋がる社会的環境管理能力形

成のあり方について考察する。

表3のシステム形成期で、行政は自身の能力形成とともに、市民の環境意識の醸成、企業への社会的圧力の形成の役割が指摘されているが、必ずしも行政自身の能力形成として、市民および企業の能力形成を促す能力要素が形成されるとは限らない。つまり、Gcapは向上するが、Gcap Fcap およびGcap Ccapの矢印の部分である、相互関係に基づく他アクターの能力形成の促進が上手く機能しない状況も想定される。これより、能力形成を進めていくにあたり、アクター間の相互関係を踏まえ、他アクターの能力形成を促すようなアクターの能力形成が求められる。

したがって、Gcap FcapおよびCcap、という社会的環境管理能力の形成プロセスを踏まえて、表1のマトリクスを用いて各アクターの能力水準を3つの能力要素から把握し、表3のステージ論での現在の位置づけを明確にすることが求められる。つまり、Gcapの水準、Gcap FcapおよびCcapにおける他アクターの能力形成促進に係る能力要素の水準、FcapおよびCcapの水準に関して、いずれの状況がボトルネックかを明らかにする必要がある。これにより、大気汚染対策の効率的・効果的实施に向けて、能力形成の対象とすべきアクター選択やステージ論に基づく具体的な達成目標、形成対象となる具体的な能力要素などを示すことができる。その評価結果は、図12のようなイメージで示される。

図12 アクター・ファクター・マトリクスに基づく社会的環境管理能力の評価イメージ



(出所)筆者作成

5. 結論

本研究は、日本の3都市の1971～2000年における大気汚染対策を事例に、アクター間の役割および各アクターの環境管理能力に一定の代替性があるとの仮説に基づき、大気汚染対策に係る社会的環境管理能力の形成プロセス、および各アクターの環境管理能力の形成プロセスに係る因果関係を実証的に示した。結果、以下のことを明らかにした。

(1) 大気汚染対策に係る社会的環境管理能力および各アクターの環境管理能力の推移を踏まえ、Stepwise Chow Testにより、社会的環境管理能力の構造変化の時期を示した。そして、構造変化前後の期間それぞれでのアクター間の相関関係を考察し、構造変化後の負符号の相関係数に基づき、アクター間の環境管理能力に、一定の代替関係が生じていることを明らかにした。これより、社会的環境管理能力は、能力水準に応じたアクター間の役割の変化・代替を経ながら形成されていくという能力形成プロセスの全体像を都市ごとに示した。

(2) アクター間の代替が起こるまでの期間を対象に、「行政の環境管理能力の向上が、市民および企業の環境管理能力を向上させるといった社会的環境管理能力の形成プロセス」に係る因果関係を、構造方程式モデルにより実証的に示した。

(3) 大気汚染対策に係る社会的環境管理能力の形成プロセスを踏まえ、アクター・ファクター・マトリクスにより各アクターの能力水準を把握し、ステージ論での現在の位置づけを明確にすることで、能力形成の対象とすべきアクター選択やステージ論に基づく達成目標、形成対象となる能力要素などを示すことができる。

謝辞

本研究は、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。また、匿名の査読者の方々から詳細にわたり大変貴重なご意見を頂きました。深く感謝の意を表します。

注記

- 1 今村(2002)は、「より良いガバナンスを生み出す」というとき、繰り返し問われることになるであろう問題の一つが、参加主体の能力問題である。政治家も、行政職員も、事業者も、そして市民も、そこに参画する主体すべてが、その能力や資質を問われることになる。」と、ガバナンスでの各アクターの能力向上の必要性を指摘している。
- 2 GRは、量だけでなく、質の把握が重要となるため、Barro(1991)らが、労働力の質の計測で教育年数や労働者所得等を代理変数で用いていることを踏まえ、環境政策に携わる人員の質を、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、組織力として示した。
- 3 村上・松岡(2006b)の図4(P49)で示したように、環境問題に関する知識・情報が定期的に得られる。その図4での「環境問題」関連図書は、NDL-OPAC(国立国会図書館蔵書検索システム)での、「環境問題」をタイトルに含む和図書の出版年ごとの検索結果件数である。これと、本研究の横浜市、名古屋市、大阪市のCKとの相関係数は、0.605、0.358、0.762となり、一定の相関が得られる。これより、CKは、環境意識、環境への責任感の水準等を測るものとなる。
- 4 社会教育に限らず、行政サービスの提供量(予算配分)は、規範的には、市民ニーズを充足させる量を過不足なく提供するものである。したがって、施設運営等に係る予算は、各施設の利用状況、および諸集会や学級・講座、講習会への参加等の利用実績、また市民の要望に基づき定まるといえる。したがって、社会教育費は、市民の環境等に関する知識・情報の交換・共有に係る「場」の活性化の程度に近似すると捉えることができる。もちろん、現実には、行政サービスの効率性や、市の財政事情により歪みが生じている可能性もある。なお、社会教育費の支出項目別内訳(平成16年度・全国)としての、消費的支出、資本的支出、債務償還費の割合は、56.5%、17.2%、26.3%となっている。

- 5 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。
- 6 社会的環境管理能力の発展ステージ・モデルは、井村・勝原（1995）、原嶋・森田（1995）、李（1999）などの先行研究におけるステージ論の検討をふまえたものである（松岡他2004）。
- 7 相関係数はスピアマンの順位相関による。
- 8 モデルが識別されない状態とは、求める自由母数と方程式の数の関係で、想定するモデルでの解が一つに定まらない状態である。そのため、モデルを識別させるために、パス係数や分散を固定させるなどの制約を課す必要がある。一般的に、潜在変数の分散を1、潜在変数から観測変数のパス係数の一つを1に固定する制約方法が用いられる。

参考文献

- 今村都南雄（2002）「公共空間の再編」、今村都南雄編『日本の政府体系』、成文堂、1-22
- 井村秀文・勝原健編（1995）『中国の環境問題』、東洋経済新報社
- 小塩隆士（2002）『教育の経済分析』、日本評論社
- 狩野裕・三浦麻子（2002）『AMOS、EQS、CALISによるグラフィカル多変量解析』、現代数学社
- 金原達夫・金子慎治（2005）『環境経営の分析』、白桃書房
- 小泉秀樹・西浦定継編（2003）『スマートグロース』、学芸出版社
- 紺野登・野中郁次郎（1995）『知力経営』、日本経済新聞社
- 高林喜久生（1988）『日本経済のマクロパフォーマンス』、東洋経済新報社
- 二宮正司（1977）「Stepwise Chow Test」、『季刊理論経済学』、28（1）：50-60
- 野中郁次郎・泉田裕彦・永田晃也（2003）『知識国家論序説』、東洋経済新報社
- 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
- 廣松毅（1986）「情報化指標の試みと「情報流通センサス」」、林周二・中村隆英編『日本経済と経済統計』、東京大学出版、97-115
- 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・岡田紗更・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
- 松下和夫（2002）『環境ガバナンス』、岩波書店
- 村上一真・松岡俊二（2006a）「都市大気汚染政策における社会的能力の評価」、『日本評価研究』、6（1）：55-69
- 村上一真・松岡俊二（2006b）「社会的環境管理能力の評価手法に関する研究：都市大気汚染対策を事例として」、『日本評価研究』、6（2）：43-58
- 村上一真・松岡俊二（2006c）「都市大気質と社会的環境管理能力の因果関係の分析」、『環境情報科学論文集』、20：309-314
- 村上一真・松岡俊二（2006d）「都市大気質と経済成長および社会的環境管理能力の因果構造分析」、『HICEC Discussion Paper』、2006-1
- 山本拓・Zhai Gou-hao（1995）「日本経済の構造変化」、倉澤資成・若杉隆平・浅子和美編『構造変化と企業行動』、日本評論社、351-365
- 李志東（1999）『中国の環境保護システム』、東洋経済新報社
- Barro, R. J. (1991). Economic Growth in a Cross Section of Countries, *Quarterly Journal of Economics*, 106, 407-443.
- Stern, P. C. (2000). Toward A Coherent Theory of Environmentally Significant Behavior, *Journal of social issues*, 56 (3), 407-424.
- Takeuchi, Y. (1991). Trends and Structural Changes in Macroeconomic Time Series, *Journal of the Japan Statistical Society*, 21, 13-25.

(2007.1.30受理)

A Causal Analysis for the Development Process of Social Capacity for Environmental Management: The Case of Urban Air Quality Management

Kazuma Murakami

Shunji Matsuoka

Tatsuo Kimbara

Graduate School for International Development and Cooperation,

Hiroshima University

mkazuma@hiroshima-u.ac.jp smatsu@hiroshima-u.ac.jp t-kimbara@hiroshima-u.ac.jp

Abstract

This study examines the causality of the development process of social capacity for environmental management (SCEM), which contributes to urban air quality improvement realized by social actors (i.e., government, firms, and citizens) and factors (i.e., policy & measure, resource management, and knowledge & technology). By applying the stepwise chow test and correlation analysis to a panel data set for years between 1971 and 2000, we show the substitutability of actors' capacity in the case of 3 major cities (i.e., Yokohama, Nagoya, Osaka) in Japan as the development process of SCEM. Then, a structural equation model is applied to examine the causality of the development process of SCEM for the period before the substitutability of actors' capacity. The model shows that increasing government capacity has increased firms' and citizens' capacity.

Keywords

Social Capacity for Environmental Management (SCEM),
Development Process of SCEM,
Structural Equation Modeling, Air Quality Management

【研究論文】

政策評価の客観性をめぐる一考察 - 不良債権処理に伴う雇用対策評価をめぐって -

西本 哲也

同志社大学大学院総合政策科学研究科

ebg1012@mail3.doshisha.ac.jp

要 約

不良債権処理に伴う失業者に対する雇用対策として、緊急雇用対策交付金の創設を中心としてセーフティネットの構築が図られ、2001年度補正により予算も付けられた。しかし、この雇用対策が本当に効果的であったかどうかは、実はよくわかっていない。雇用対策の実施を担った厚生労働省は「おおむね効果があった」と自己評価していたが、報道や国会質問により、自己評価に対する疑問や政策の効果に対する疑問が呈されていたためである。

他方、わが国の政策評価制度は、「評価の評価」(メタ評価)を行なう機能があらかじめ設定されている。その役割を担うのは、総務省行政評価局による客観性担保評価と呼ばれる活動である。この緊急雇用対策交付金に対しては、客観性担保評価が一歩進んだ形で行なわれたが、同時に課題や限界も明らかになった。

キーワード

メタ評価、客観性担保評価、行政評価局、緊急雇用対策交付金

はじめに

わが国の政策評価制度は、各府省が所管する政策について自らが評価(以後「自己評価」)を行ない、企画・立案に反映させることが基本原則のひとつになっている。したがって、自己評価が「お手盛り」評価に陥るといふ危惧は避けられず、これを払拭するため、評価の客観性を高める仕組みがあらかじめ設定されている。その代表が、総務省行政評価局による「評価の評価」を行なう客観性担保評価の活動であり、これは「行政機関が行なう評価に関する法律」(2001年6月成立)(以

下「評価法」)により規定されている。

客観性担保評価の取り組みは、総務省行政評価局により毎年報告されているが、客観性担保評価の現状とその限界を示す事例が出現した。それは、小泉構造改革の不良債権処理に伴う一連の雇用対策への評価活動である。雇用対策を実施した厚生労働省は、「おおむね効果があった」と自己評価していたが(2002年度厚生労働省実績評価書)、その認識とは大きなズレのある実態が報道機関によって指摘されていた。これに応じて総務省も厚生労働省に対する客観性担保の活動を行ない、ある意味において客観性担保の活動が一歩進んだ形

で実施されたのである。

そこで以下では、この緊急雇用対策に対する一連の評価活動を事例研究としてとりあげ、わが国の客観性担保評価について、次の3点を明らかにしていく。1点は客観性担保評価の「制度としての」役割や位置づけの紹介である。そして2点目が、自己評価に対して客観性担保評価が「どのように機能しているのか」という実態を明らかにすることである。3点目は、先の2点を通して、現状での客観性担保評価の課題や限界を明らかにすることである。

1. わが国の政策評価制度における客観性担保の仕組み

ここでは、「制度として」の客観性担保評価の中心的な役割を担う総務省の客観性担保評価の役割や位置づけの紹介を行なう。もちろん、評価の客観性を担保する活動は、総務省の客観性担保評価のみによって担われているわけではない。そこで、評価制度「全体としての」客観性担保の仕組みを記述し、次に総務省の客観性担保に期待されている役割や位置づけを確認した上で、この客観性担保評価の内容について具体的に説明していく。

(1) 政策評価制度の概略

政策評価制度が公式に盛り込まれた最初の規定は、中央省庁等改革基本法（1998年6月成立）である。その後の「中央省庁等改革の推進に関する基本方針」（以下「基本方針」）（1999年4月中央省庁等改革推進本部決定）では評価法の基本的な枠組みを構築するため、標準的ガイドラインを策定することが指示された。それを受け、政策評価の手法等に関する研究会を中心として検討が行なわれ、「政策評価に関する標準的ガイドライン」（以下「ガイドライン」）（2001年1月15日、政策評価各府省連絡会議了承）が策定された。その後、政策評価制度の法制化に関する研究会が報告・取りまとめを行ない、ガイドラインを踏襲した評価法が成立する運びとなった。

評価法の基本的な原則は以下の通りである。まず、評価法は国民に対する説明責任の徹底、国民本位の効率的で質の高い行政の実現、国民的視点に立った成果重視の行政への転換という3点を目的としている。また、各府省が所管する政策について、その効果を把握し、必要性、効率性、有効性などの観点から評価を行ない、政策の企画立案や政策に基づく活動を的確に行なうために重要な情報を提供するということが定められている。さらに、評価制度は基本的に全政府的に導入されており、各府省独自に基本計画や実施計画の策定や評価結果の公表が行なわれている。

(2) 客観性担保の仕組みと総務省の役割

評価の基本枠組み

政策評価の客観性を議論する際に重要な要素となる「評価主体」についての考え方は、「行政改革会議最終報告」からある程度固まっていた。すなわち基本枠組みは、各府省が自己評価をすることが大前提であるが、全政府的を見渡す立場から、「評価を評価する」役割を総務省に与えるというものである。

各府省内で実際の評価の中心を担うことになるのは主に各「課」である。各課で行なう評価の結果をPLAN（企画立案）DO（実施）SEE（評価）のマネジメント・サイクルに反映させるためである。

各府省内にはさらに評価を専門的に担う評価専任組織（たとえば政策評価課あるいは政策評価官、または政策評価室）が置かれている。この評価専任組織は、府省内の評価の厳正かつ客観的な実施を確保するために、評価計画や基本事項を作成し、評価に対する技術的な助言や府省内全体の評価の取りまとめや公表などの総括事務を担っている。

政策評価制度における客観性を担保するための仕組み

自己評価を基本とするにはそれなりに理由があるが、恣意的な評価に陥りやすく、またいわゆる「お手盛り評価」になってしまうなどの危惧が常に付きまとう²。当然ながら、評価の客観性を高める必要性から、いくつかの仕組みがあらかじめ

設けられている。評価の客観性担保に係る基本的な考え方や仕組みは以下の通りである。

まず、客観性を高めるための基本的な指針のひとつとして、「厳格な手続きの執行」という考え方があり、基本方針では、評価の手続きを厳格に行なう方針が明確に打ち出され、評価の枠組み（評価計画・評価システムの設計）の手順の妥当性、データの信頼性・妥当性、評価結果とその根拠の整合性・妥当性を検証可能にすることで客観性が高められるとの考えが記述されている。ここでは、法的な手続きや会計基準などの各種規定が遵守されているのが評価の基準とされ、いわゆる監査的な発想がもたれている³。

次に、各府省における第三者の活用があげられる。高度な専門性が必要である場合や多様な意見を反映させる必要がある場合、もしくは政策の性質や内容によって必要な場合に行なわれ、学識経験者や民間の有識者等から個別に、もしくは組織を設置して、意見の反映が行なわれている。

また、評価結果やその過程をHPなどにより公表することも、客観性担保にとって重要な方法のひとつである。評価の結論だけでなく、評価の際に使用した仮定等の前提条件、評価手法・指標、データ、学識経験者の意見等の過程を含めて、可能な限り具体的に公表することが評価法や基本方針により規定されている。

さらに、国民の側からの意見を反映させることも可能である。基本方針では、各府省が国民からの評価に対する意見や要望を募集する窓口を設置することが規定されている。他方、後述する総務省の客観性担保評価の役割にも関係するが、総務省は独自に情報収集を行っており、各府省が実施した政策評価について国民が寄せる情報を受け

付けている。具体的な対象は、「実態を反映していない」「つじつまが合わない」などの問題がある場合、さらに、「各府省が政策評価を実施しているかどうか分からないが、国の施策・事業について、『効果があがっていない』『社会状況の変化により、必要性が低下している』などの理由により、中止や見直しが必要ではないか疑問がある場合」などを例示している。また、本稿で取り上げるように、報道機関による評価に対する報道も情報収集の範囲である。

(3) 総務省の客観性担保評価の活動内容

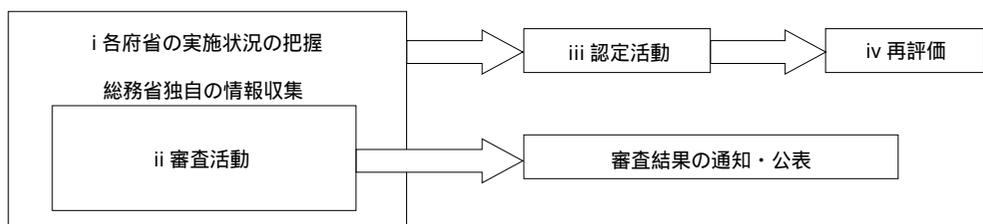
これら評価の客観性を確保する仕組みの中核を担うのが総務省行政評価局による客観性担保評価である。客観性担保評価とは各府省が行なう政策評価について、各府省の政策の見直し、さらには改善の推進を図るために総務省が「評価の評価」を行なう活動である。

具体的な活動としては、形式的な「手続き」にかかわる審査活動、評価「内容」が適切であるかを吟味する認定活動、不十分である場合には総務省が各府省に代わって直接評価を行なう再評価の活動として設定されている⁴。それぞれの活動は互に関連しており、認定活動は審査活動を中心とした実施状況の把握や独自の情報収集により認定活動を行ない、再評価は認定活動を前提として実施される⁵（図1参照）。

審査活動（図1のii）

審査活動は、正式には「各行政機関が実施した政策評価について、その実施手続き等の評価の実施形式において確保されるべき客観性・厳格性の

図1 客観性担保評価の活動の流れ



（出所）総務省行政評価局パンフレット（2004）総務省行政評価局

達成水準等に対する審査」(基本方針)と呼ばれる活動である。審査活動の目的は、各府省が行なった自己評価の達成水準等に関して、法定要件などの点検項目を中心に「手続き」が守られてかどうかを審査し、今後の課題を明らかにすることである。

その形式としては、個別の審査と総括審査がある。個別審査とは各府省の個別の自己評価について、評価の枠組み(計画・設計)に係わる手順等の網羅性・充足性に関する項目、評価に使用したデータ・資料等の信頼性に関する項目、評価結果とその根拠の整合性に関する項目を基本として、主に手続き面での点検を行なうものである。他方、総括審査とはより大枠で、実績評価方式、事業評価方式、総合評価方式といった「評価方式」を軸とし、政策評価制度の全体におよぶ課題等を明らかにするものであり、年次報告と言うこともできる。

審査活動は客観性担保評価の中でも活発に行なわれており、個別審査と総括審査は共に報告書やHPを通して多くの成果を見ることができている。

認定活動(図1の iii)

認定活動とは、正式には「再評価等の実施の必要性の認定活動」と呼ばれる。個々の政策評価について疑問が生じた場合や社会情勢が変化したにもかかわらず各府省が評価を行っていない場合に、各府省に対して事実関係の照会を行ない、それでも疑問が解消しなかった場合は総務省が政策評価・独立行政評価法人評価委員会の審議を経て、評価のやり直しなどの認定を行なう活動である。つまり総務省が各府省の行なった評価の「内容」に関して疑問を呈す活動であり、具体的には、評価の適切さ(たとえば、適切な指標を設定しているのか)や必要性(たとえば、問題がある政策なのに評価が実施されていない)などを対象として、この活動が行なわれる。しかしその活動は、2006年3月行政評価局の『政策評価の点検結果評価の実効性の向上に向けて』で、ようやく認定に向けた活動が報告された程度である。

再評価(図1の iv)

再評価とは、認定活動の結果、当該行政機関に

ゆだねては評価の客観的かつ厳格な実施が確保されないと認定されたときに総務省が実施する評価のことであり、それも総務省が単独で各府省に代わって評価を行なうものである。正式には、「客観的かつ厳格な実施を担保するための評価」と呼ばれ、客観性担保評価のねらいや目的から考えれば中心となるべき活動であろう。

しかしながら、再評価に関する活動は、認定活動も十分に行なわれていない中、いまだ実施されていない。総務省の検討も認定の充実までになっており、再評価がいつどのように実施されるかは不透明である⁶⁾。

2. 不良債権処理に伴う雇用対策と自己評価

次に、「制度としての」客観性担保評価が現在「どのように機能しているのか」という実態を明らかにするため、小泉構造改革の不良債権処理に伴う雇用対策の評価をめぐる厚生労働省と総務省の間の一連の活動を事例研究としてとりあげたい。

(1) 雇用対策の背景とその内容

本稿で取り上げる雇用対策は、第一次小泉内閣の目玉であった不良債権処理問題に伴って生み出される失業者へのセーフティネットを緊急的に構築するための事業とし考案されたものである。不良債権の最終処理の過程では企業の不採算部門の整理・縮小を伴い、企業倒産につながる可能性もあることから、雇用への影響が懸念され、雇用セーフティネット拡充に繋がっていった⁷⁾。

これらの方針を策定し、具体化させる作業は、経済財政諮問会議主導で立案された。不良債権処理を中心とした構造改革全体の基本方針がはじめに示されたのは、「今後の経済財政運営及び経済社会の構造改革に関する基本方針」(いわゆる第1弾骨太の方針)である。第1弾骨太の方針では同時に、雇用に対する配慮もなされており、「不良債権処理の影響に備えたセーフティネットの拡充」、「雇用への影響を最小限に抑えるため、雇用対策法、雇用保険法、離転職者向け教育訓練、緊

急雇用創出特別奨励金等の制度・施策を活用」を図ることも明記されていた。また、この方針を具体化するため、期限や数値具体的な事業内容を盛り込んだ「改革工程表」(2001年9月26日)が策定され、さらに骨太の改革と関連付けられた経済対策として「改革先行プログラム」(2001年10月26日)、「緊急対応プログラム」(2001年12月14日)の2つのプログラムも策定されていた。

これらの中で、雇用対策の目玉として位置付けられたのが、新公共サービスによる緊急かつ臨時的な雇用・就業機会の創出という考え方である。新公共サービスとは、具体的には、学校の補助教員として「学校いきいきプランに基づき多用な経歴を有する社会人を教員補助者として学校に受け入れ、教育活動を充実する」、警察支援要員として「防犯指導、駐車違反對策等を強化する」、環境保全のための森林作業員として「森林整備の強化を通じて環境保全を図る」などが想定されている。この新公共サービスの領域で「緊急かつ臨時的な雇用」の創出を図るために、2001年度補正予算では、緊急雇用対策交付金として約3,500億円の予算が付けられた。

この緊急雇用対策事業の実施を担うのは、厚生労働省から交付金を受ける地方自治体であった。地方自治体が地域のニーズを踏まえ、上記に例示したもの以外にも、独自に創意工夫を凝らした雇用の創出を図るためである。ただし、あくまでも

「緊急かつ臨時的な雇用」の創出が目的であるため、雇用期間は原則として6ヶ月までと決められていた。

(2) 雇用対策に対する厚生労働省の自己評価⁸⁾

評価の体系と体制

厚生労働省の評価は、基本目標 - 施策目標 - 実績目標の体系で評価を行なわれている。評価の対象は、施策目標と実績目標であるが、実績目標には、達成状況を測定するための評価指標が設けられているものの、施策目標に関する評価指標の設定は厚生労働省の基本方針によれば努力目標である。「実績目標」では、設定された「評価指標の達成の度合い」と「有効性」「効率性」の観点から自己評価が行なわれている。これらの自己評価を踏まえて、「反映方針」という形で自らの企画・立案に活かされることになっている⁹⁾。

緊急地域雇用対策交付金に対する実績評価の概要

厚生労働省は、先の緊急雇用交付金への実績目標として「緊急地域雇用創出特別交付金を活用して、各地域のニーズを踏まえた事業を実施し、公的サービス部門において緊急かつ臨時的な雇用・就業機会の創出を図ること」を掲げている。評価指標としては、「交付金の支出額」と「新規雇用・就業者数」を設定している(表1参照)。

表1 厚生労働省実績評価の概要

評価指標	有効性の評価	効率性の評価	総合的な評価
事業支出額 目標：139,600 百万円 実数：133,900 百万円	地域ニーズに応じた臨時・応急的な雇用創出が図られており、有効であったと考えられる。また、半年後の就業状況調査によると、平成14年9月30日現在で就職している者が57%にのぼっており、臨時的な雇用機会を提供することによって、就業者にとって再就職に至る「つなぎ」として一定の機能を果たしていると考えられる。	都道府県が事業主体となることにより、地域の実情に応じた雇用・就業機会の創出が効果的に図られている。また、各都道府県からの要望をふまえ、平成14年12月より、事業額に占める人件費割合、雇用期間等について運用改善を実施し、より効果的な事業の実施を可能としたところである。	地域雇用開発促進助成金、地域求職活動援助事業、緊急地域雇用創出特別交付金の活用により、地域の実情に即した雇用機会の創出等が図られ、目標をほぼ達成したと考えられる。
新規雇用・就業者数 目標：140,000 人 実数：185,000 人			

(出所) 2002年厚生労働省実績評価書
<http://www.mhlw.go.jp/wp/seisaku/jigyoyou/02jisiseki/4-2-2.html> (2006年11月)

データの集計は、地方自治体からの報告書である「緊急地域雇用創出特別基金事業実績報告書(2001年度第4半期)」によって行なわれている。緊急雇用創出特別事業実績報告書は、緊急雇用創出特別基金事業実施要綱に基づいて都道府県からの報告がまとめられている。事業の実績を把握することを目的として事業額、新規雇用された失業者数、新規雇用された失業者のうち正式雇用した者の数などを把握している。

報告書の集計の結果、設定した2つの評価指標は、「支出金額」においては厚生労働省が設定した目標をほぼ達成しており、「雇用者数」は目標よりも上回っていることがわかる。さらに、有効性の評価の欄では、緊急的・臨時的雇用の創出を目指すという目的が達せられたかどうかについての厚生労働省の認識が象徴的に現れている。すなわち「臨時的な雇用機会を提供することによって、就業者にとって再就職に至る『つなぎ』効果(事業終了後、安定した雇用につながっているかどうか)として一定の機能を果たしたと考えている」との厚生労働省の認識である。

自己評価に見られる問題点

この緊急雇用対策特別交付金に対する実績評価の概要を見ると、自己評価に対するいくつかの象徴的な課題や問題点が指摘できる。

1点目は、評価指標が就業人数や支出金額などアウトプット(生産物)レベルになっており、アウトカム(成果)やインパクト(波及効果や副次的効果)レベルで効果を測定していない。緊急雇用創出特別奨励金において雇用された人は、次の就職が見つかるまでの「つなぎ」としての効果はあったが、継続的な雇用に結びついているのかどうか、さらには「つなぎ」として一時的に雇用されたからといって、そもそもの目的であるセーフティネットとしての役割をどこまで果たしたのかという部分は評価書からはわからなかった。

2点目は、「骨太の方針」や「緊急対策プログラム」という政策レベルで重要な指針が契機となった重要度や緊急性が高い政策であっても実績評価の段階になると重要方針からのつながりが見えず、評価においてその他の既存の施策・事業との重要さや緊急性の点で違いが感じられない。政府

が重要な位置づけを与えている施策・事業も、縮小していくことが明らかな施策・事業も評価書の位置づけでは同様の扱いであった。

3点目に、今回新設や大幅な拡充が行なわれた雇用対策以外にも、同様の目的を持った助成金はすでに存在しており、新旧の政策の関係が見えないことである。一連の雇用対策によって成果が出た(たとえば、「雇用が生まれた」、「失業率が改善した」と)との認識が示されているが、既存の施策や事業の効果や好・不況などマクロ経済の影響がどれくらいあるのかが全く考慮されていない。

以上に列挙した問題点は、各府省に設置する第三者機関や総務省の客観性担保評価のプロセスを通じて洗い出され、各府省のマネジメント・サイクルに活かされ実務が改善されるのならば大いに意味があるはずである¹⁰。また、そのためには長期的に評価の質的改善を目指す視点も必要になるであろうが、明らかに問題のある評価であると考えられる場合には、たとえば「再評価」に期待されるような、是正に向けた力を働かせていくことも客観性担保評価の役割として必要になるだろう。そこで次に、この雇用対策を事例として、評価への問題提起や是正に向けた客観性担保の活動がどのように行なわれたのかを検討する。

3. 雇用対策の自己評価に対する客観性担保評価の活動

緊急雇用対策に対する厚生労働省の自己評価と、総務省の客観性担保評価活動の一連のやり取りを見ることで、制度としての客観性担保評価(とくに認定から再評価の内容面での客観性担保評価の活動)が現在「どのように機能しているのか」、その実態を理解することができる。この雇用対策に対する客観性担保評価の活動は、2つの特徴を持っているからである。1点目は、認定活動に向けた「照会」が早くから(評価法実施の2年目である2003年度から)行なわれたからである。2006年3月総務省行政評価局の『政策評価の点検結果』によって認定活動に向けた照会の活動が報告されているが、それらに先立って行なわれた活動であり、どのような見直しが行なわれたのか見

ることができ、当面の課題を知る上でも有益である¹¹。2点目は、報道機関が雇用対策への問題提起を行ない政策に対する疑問が広がっていたためである。報道機関が事業やその評価に対する疑問を指摘し、これが総務省の客観性担保評価活動のきっかけとなったためである。また報道機関と同様の指摘が国会質問で行なわれるなど、認定活動が行ないやすい環境にあった。

(1) 報道機関による雇用対策とその評価への問題提起

不良債権処理問題に関する雇用対策に対する厚生労働省の自己評価は、「設定した評価指標を達成し、一定の効果을上げており有効性がある」という認識であった。しかし、報道機関を中心に異なった見解が示されていた。

指摘の中心となったのは朝日新聞の「雇用対策総点検」と名付けられた連載である。連載は2003年10月13日から2004年3月30日まで定期的に行なわれ、その中で「雇用対策が成功していない」「数々の問題がある」などの現場の声を紹介してきた。記者による取材（受給者や事業者へのインタビュー）や専門家（雇用や労働問題に詳しい大学教授、コンサルタント、社会保険労務士など）の指摘、さらには評価書や決算書などから情報を得て雇用対策やその評価に対する問題点の指摘を行なっている。指摘の内容は次の4点に整理できる。

雇用対策に対する厚生労働省の自己評価に関する問題点の指摘

まず、評価指標の選定に関して厚生労働省は、予算の消化率、助成金の支給実績や支給決定人数といった指標だけを取り上げていたが、記事では仕事そのものが有益であったのか、失業者がその後どうなったのか（たとえば継続的に雇用されているのか）で効果を計るべきであると指摘している¹²。

この点に関しては、坂口厚生労働大臣（当時）が経済財政諮問会議において、アウトカムを含めた数値目標を立てる必要がある、と説明するまでに至った¹³。さらにこの点は国会においても質

問がなされ、短期間の臨時雇用では一時しのぎにしかないとして、国会でも「つなぎの臨時雇用だとしても、常時雇用の道筋をつけないと役に立たない」という趣旨の質問がなされた¹⁴。

自己評価のデータの信頼性に対する指摘

さらに、緊急雇用対策の評価を行なうデータの基になる調査対象の抽出が、よい結果が出るように意図的に決められたのではないかという疑問も記事の中で提起された。厚生労働省の評価書では「半年間の交付金事業終了後に就職しているものが57%である」と説明されていた。評価書公表後に行った追跡調査によると、その内訳は、「雇用された事業所とは別の新しい仕事に就いた」（42%）と継続雇用（13%）起業したもの（2%）であると分けて示されていた。しかし、朝日新聞が追跡調査をしたデータによると、その事業所に継続して雇用された人は1.6%に過ぎなかった。よい結果が出るように、継続雇用をした事業所を意図的に抽出し、対象から外していたという疑念が提示されたのである¹⁵。

助成金による雇用対策が成功していないとの指摘

さらには、緊急雇用対策だけの問題ではないが、雇用対策が全体として成功していないという指摘も記事の中でなされている。その根拠としては、使われていない助成金が多数あるとの指摘であり、1999年から2003年までの10の助成金・奨励金の予算と決算を調べ、消化率がトータルで50.7%であったという事実を挙げている¹⁶。それだけではなく、助成金に関連して不正受給が相次いでいるという事実も示している。受給者となる事業者がペーパーカンパニーを設立して不正受給を受ける、労働局職員による裏金作りや着服など汚職、事業主のカラ出向問題などさまざまな形で不正や汚職が雇用対策に関係する助成金をめぐって起こっていたためである。さらに会計検査院の調査の結果、同一人物で助成金を二重取りしているケースが全体の1割にのぼり、総額で51億円もの誤った支給が行なわれていた¹⁷。

雇用対策の制度上の問題点の指摘

記事の中では、制度上の問題点まで視点を広げ

た指摘や提案も述べられている。その代表が、助成金という手段に偏重しているという指摘である。「助成金を出すから人を雇え」という姿勢を批判し、いわゆるばらまき型から成果型への転換を図るべきだとの主張が述べられている¹⁸。

また、今回の緊急雇用交付金は確かに地方自治体が実施主体として創意工夫を行なうことで受給者数や就労者人数を確保することができたが、本来の不良債権処理によって生まれる失業者に対する対策ではなくなっている場合も多く見受けられるという。たとえば、自治体によっては高齢者の就労支援（シルバー人材センターへの委託）や障害者の雇用対策（たとえば、障害者地域作業所運営指導事業）、ホームレス対策など「不良債権処理に伴う」失業者ではない対象者に多くの助成金を使っている場合が見受けられた。とくに大阪市におけるホームレス対策（あいりん日雇い労働者等事業）には、平成15年度緊急地域雇用創出特別基金事業実績によると、新規就労者数が大阪市の事業の中で、約50%に上っていた¹⁹。

緊急雇用交付金は2005年3月31日をもって終了したが、このような新聞報道を見る限り、今後緊急的な雇用を創出するという場合に、自治体の裁量に任せるといった手段が適切であったのか否ということには疑問が残る。

(2) 総務省が行なった客観性担保評価の活動

雇用対策に対してこのような批判や疑問が聞かれる中、総務省は認定活動に向けた照会の活動を行った。照会は認定の前段階の活動として、評価に疑問を抱く、もしくは問題点が明らかになると総務省が考える場合に、まず各府省に対して事実関係を問い合わせ、回答を求める活動である。緊急地域雇用創出特別交付金に対しては、2つの疑問点について照会が行なわれた²⁰。

政策効果（「つなぎ」効果）についての指摘

まず、第1点目は、「つなぎ」効果（事業終了後、安定した雇用につながっているかどうか）も評価の対象になるのではないかとこの疑問からの照会である。そこで、「つなぎ」効果も評価指標として認識しているかどうか、またそれも踏まえて当該

交付金が効果を上げているか、また今後事業を継続していくか否かの判断をするのか等について厚生労働省に回答を求めた。照会の根拠としては、評価法第3条の把握すべき政策効果に関する記述を挙げている。評価法第3条によると把握すべき政策効果とは、「当該政策に基づき実施し、又は実施しようとしている行政上の一連の行為が国民生活および社会経済に及ぼし、または及ぼすことが見込まれる影響」のことである。

これに対する厚生労働省の回答は、「つなぎ」効果はあくまで副次的効果であり、評価指標とはならず、政策効果には該当しないというものであった。

この回答を踏まえた総務省の見解は、以下の2点である。第一に、緊急地域雇用創出特別交付金の目標が達成されたかどうかは、「つなぎ」の効果の有無にかかわらず、交付金の活用により「新規雇用・就業」が創出されたかどうかで判断すべきであるということ。第二に、政策の意図が交付金の活用による「新規雇用・就業」の創出である場合には、今回行なわれた評価において把握すべき効果が把握されていなかったことにはならず、したがって、効果の把握範囲について評価法上の問題は生じないということである。

評価に使用したデータの信頼性についての指摘

次に、緊急地域雇用創出特別交付金の活用に関する就業状況調査のデータの信頼性に対する照会活動である。就業調査に関する指摘は2点あり、1つ目は調査方法に関する内容であり、報道機関によって就業状況調査が複数の県で無作為抽出されていないとの指摘を受けて、厚生労働省の認識を問う内容である。総務省はデータの信頼性に疑念が持たれかねない評価であると考え、調査方法についての回答を厚生労働省に求めた。2つ目は、新たに全都道府県からの報告を求めたデータによると、継続雇用は1.6%という数字になっており、どのような目的のもとに別の調査が行なわれ、結果としてなぜ著しく異なった数値が表れたのかということに関する照会である。

1つ目の調査方法に関して厚生労働省の回答は、都道府県へのサンプル抽出の指示が文章上不十分であったこと認めつつも、継続雇用者が増えるよ

うに調査対象者のサンプル抽出を行なう意図を都道府県が持っていたわけではなく、その方法は「おおむね無作為抽出により」行なわれていたと回答している。同時に厚生労働省は、恣意的な調査をおこなったのではないかとの疑念が生じないように、サンプル抽出の方法の指示を適時に行なうという認識も示している。

2つ目のデータの食い違いに対する厚生労働省の回答は次の通りである。評価のデータは都道府県からの報告にもとづいた「緊急地域雇用創出特別基金事業実績報告書（以下「実績報告書」）」をベースに作られているが、他にも「緊急地域雇用創出特別基金事業従事経験者の就労状況に関する調査」（以下半年後の「就労調査」）として実態の調査が行なわれていた。朝日新聞が着目した1.6%というデータは、実績報告書に基づいた数字で「引き続き新規雇用された事業所で正式雇用された」者の数（23,000人のうち362人）であり、実績報告書で示された57%というデータは、交付金事業で新たに雇用された者の「現在の就業状況」で半年後の就労調査に基づいている¹⁰。つまり、事業所（新公共サービスの実施主体となる都道府県の出先機関やNPOなど）で継続して新規雇用をされた失業者は1.6%で、就労後半年でなんらかの形で就労している人は57%であったとの回答であった²¹。

この回答に対して総務省は、それぞれの目的に沿って把握しようとする対象の違いであると理解を示し、厚生労働省に対して今後結果として疑念が生じないように、サンプル抽出依頼について改善していくとの確認を行なうにとどまっている。

（3）総務省による客観性担保評価の限界

総務省の認定に向けた照会活動は、認定や再評価を目指した一歩進んだ形であり、評価に関する疑問点を明らかにできたという意味で客観性担保評価の成果であると言えるだろう。しかしながら、一連のやりとりを見ると、「制度としての」客観性担保評価の運用が実態として機能しているとは言い難い事実も浮かび出ている。

その事実のひとつは、評価の見直しに関する限界である。今回の緊急雇用対策に対する客観性担

保の活動は、翌年以降の評価の見直しにはつながっていない。2002年度の評価書には、政策手段の有効性評価の欄においてつなぎ効果に対する記述があり、有効性の判断材料のひとつとされているが、2003年度以降の評価書においては、つなぎ効果に対する記述の部分がなくなっている。その他の部分は、数字が入れ替えられているだけで、目だった見直しは行なわれていない。効率性の項目や事業実績の項目も同様に、数字の入れ替えが行なわれているだけである²²。

また、今回の事例のように事実関係の「照会」がなされたとしても、認定に踏み込むことは困難であった。照会相手先は施策や事業に対する情報を持っており、総務省からの照会に対して認定を逃れる回答を行なうことは容易であるが、情報を持たない総務省はこの回答によって認定を行なうか否かを判断せざるを得ないため、認定には踏み込みにくいからである。

一方では、客観性担保評価の活動には、わが国の政策評価制度の本質的矛盾が象徴的に現れている。わが国の政策評価は、手続き・プロセスを細かく決め、それを厳格に遵守することを重視する「法規国家を基調とする日本の文脈」と、成果主義を目指すアングロサクソンの考え方という本質的に矛盾する2つの文脈が混在していることが指摘されている（古川 2005）。緊急雇用対策交付金の例からも分かるように、客観性担保評価は形式的な評価手順に関する審査だけでなく、評価「内容」が適切かを確認する認定に向けた活動においても、「箸の上げ下ろしまで」細かく見ていく作業を行なっている。この作業は総務省側にも各府省の側にも多大な作業負担を強いているが、政策評価が本来目指す成果志向の行政、すなわち「政策手段の選択はある程度自由で、いかに政策目標を達成できたかを問う評価」にはつながっていない。

また、客観性担保評価は、「不良債権処理に伴う失業に伴う雇用対策を緊急雇用対策事業を目玉として行う」という制度設計に対する「評価の評価」までを果たしえなかった。むしろ、当初のセーフティネットの構築という目的をどこまで達成したのか、社会的に意義が本当にあったのかどうかという見地から、事実を積み上げていくことに

よって雇用対策に対する「評価の評価」を行なったのは報道機関であった。

このように制度としての客観性担保の構造は、評価法等で記述されている「総務省による政策評価の客観的かつ厳格な実施を担保するための評価」を実現するには限界があり、政策評価制度の問題点のひとつとなっている。

(4) 客観性担保評価が機能しない理由

それでは、なぜ認定・再評価の活動が機能していないのであろうか。ここではその理由として主に次の4点が指摘されている。

第一は、平松(2005)が客観性担保評価は評価制度の設計ミスであるとまで指摘したように、認定を行なうためには総務省が評価書に記載されていない情報を知っておく必要があるにもかかわらず、実態を知るすべが用意されていない非現実的な規定になっているためである。

第二は、評価法や全府省合意(基本方針やガイドラインなど)において、「いかなる場合に客観性担保評価を行なうのか」という基準が明確でなく、抽象的な文言にとどまっていることが考えられる。評価法12条2項においては、客観性担保評価を行なう場合は、「改めて政策評価が行なわれる必要がある場合」「社会情勢の変化に的確に対応するため当該行政機関により政策評価が行なわれる必要がある場合に当該行政機関により実施が確保されないと認めるとき」と示されているだけで、どのような状況で実施が確保されないと判断するのかについては明確に示されていない。

第三に、各府省からの大きな抵抗である。各府省は手続論ないし方法論に限定した審査活動であってすらも政策内容に不当に介入していると感じており、まして政策の成否にかかわる総務省の認定や再評価に対する各府省の風当たりはかなり強い(南島2004)。認定、再評価の活動は、他の府省が行なう政策に疑問を呈するという意味で省庁セクショナリズムを壊す活動であるために、大きな抵抗感を各府省は持つ。

最後に、各府省の大きな抵抗が予想されたにもかかわらず、評価法において、総務省が客観性担保評価の活動を行なう際に必要な具体的な権限

(たとえば資料提出権や是正勧告件)が与えられていないことである。このため、報道機関や国民からの意見の募集、または各府省の見解や事実関係を確認する「照会」という方法でしか直接的な情報収集ができなかった。また評価の問題点を各府省に指摘したとしても、改善への動きは各府省の自主的な取り組みに委ねられている。加えて総務省には評価内容が適切でないすべての事業を各府省とやり合うだけの組織・人員体制や専門性も乏しいのが現状である。

(5) 客観性担保評価の限界の打開に向けた動き

このような客観性担保評価の機能不全的な状況を打開するため、総務省内では認定活動の強化に向けた検討がなされている。「『評価の実施の必要性の認定』の考え方の整理と今後の取組」(2003年8月、総務省行政評価局、以下「今後の取組み」)がその成果である。

今後の取組みでは、認定の考え方を踏まえ次の2点を重点的に取り組むべきであると記述されている。1つは認定活動の手順等の明確化であり、2つめは実際に行なわれた評価を素材として、個々の事案にどのように認定活動の考え方を当てはめていくか、その基準の明確化の重要性である。

さらに今後の取組みでは、検討対象を絞り込んでいく際の手がかりにも言及している。すなわち、内閣の基本的な方針等により重点的に取り組むべきこととされた行政分野において当該行政機関が所管する主要な政策、内外の社会情勢の変化を踏まえ見直しや改善の必要があると認められる主要な政策、国民からの評価に対するニーズが高く評価を実施する必要があると認められる政策である。これらに該当するような政策や国民からの評価の内容について何らかの疑問が出されている政策については、当該政策や評価について関係行政機関に照会し回答を求め、必要に応じてヒアリングを行なうなど関係する情報の収集・分析等の活動に幅広く取り組んでいく必要性が強調されている。

しかしながらこれらの検討は、府省間での合意事項にはなっていない。

おわりに

最後に客観性担保評価の今後について触れておく。

評価制度は一定期間経過した後に見直しを図ることが予定され（「評価法」附則、第2条）、見直しに向けての現状認識や課題の整理が政策評価・独立行政法人評価委員会による『政策評価制度に関する見直しの論点整理』（2004年3月）（以下、論点整理）や『政策評価制度に関する見直しの方向性』（以下、方向性）としてまとめられ、基本方針の改定（2005年12月16日閣議決定）が行なわれた。

論点整理や方向性の中では、客観性担保評価活動も今後の課題として取り上げられた。その中では、これまでの活動は審査が中心であったが、今後は内容面にも視野に入れた活動に一層取り組むことも示されていた。しかしながら、改定された基本方針の客観性担保評価に関する記述は、前の基本方針と変わっておらず、内容面（認定や再評価の活動）での総務省の権限（資料提出権や是正勧告権）を強化する対策は行なわれなかった。さらに、政治家も総務省の客観性担保評価機能の充実に対する認知や関心は低く、総務省の権限を強化する制度改正に向けて強力な力が働く期待も薄く、見直しの動きは弱いのが現状である。

そこで、現行の制度の枠中で認定や再評価の活動を充実させることを考えるのであれば、本稿で紹介した事例はひとつの方向性を示している。すなわち、報道機関による情報提供や指摘によって総務省の客観性担保評価が一步進んだ形で行ないえたように、報道機関、国民やその代理人である政治家が評価活動に参画することで、評価の質的向上が期待できるということである。今回の事例は、報道機関により大々的に取り上げられたことで、総務省も施策や事業に対する情報をより詳しく知ることができた。さらに国会での質問や大臣が取り組みに対して公式的な場で発言するまでに至ったことで、改善に向けた厚生労働省への圧力を強化することができ、一步進んだ形で認定に向けた活動を行うことができた。

こうした応援団を積極的に活用することには、行政府を構成する官庁のひとつである総務省が積

極的になれるかどうかという疑問もあるだろう。しかしながら、指摘してきたように、政策の適否に密接に結びつく評価「内容」に踏み込む認定や再評価という活動を行なっていくには、総務省と府省間のやり取りや調整だけでは限界があった。国民や報道機関が実態についての情報の提供を行ない、政治家がさらに評価制度に対して関心を持つことで、総務省の権限の弱さを補っていくことが、客観性担保評価の可能性を広げる鍵となるだろう。

注記

- 1 田辺国昭（2001）では、行政内部で自己評価が全政府的に、また強制的に行なわれることを指し「強制された自己評価」であると指摘している。
- 2 自己評価を原則とする理由はいくつかある。まず、行政内部で次の政策形成のサイクルに活かすという内部のマネジメントの改善への要求が強いという理由が挙げられる。また、政策に対する情報量や組織力の官・民隔差という理由も重要である。行政は日々、継続的・組織的に政策の立案や実施にかかわっており、出先機関なども含めるとかなり広範囲にわたって関わっているのに対し、民間においては情報の収集・分析の可能性に限界と制約があるためである。
- 3 厳格な手続きの執行に関する記述は、山谷清志（2002）および外務省経済協力局評価室（2003）を参照した。
- 4 正確にはこのほかにも共同評価という活動も含まれる。これは各府省と共同で再評価を行なう活動である。
- 5 総務省行政評価局（2003）によると、審査活動が必ずしも認定に直接結びつくものであるとは言えないが、審査によって得られた情報や知見は認定を行なう際にも有益な情報として活用されることが期待されている。
- 6 平松英哉（2005）によれば、総務省内で再評価は、狭義の客観性担保評価であると考えられており、俗に「伝家の宝刀」と呼ばれているようである。平松はさらに再評価の活動が行なわれていない状態を指して、再評価が抜けない「伝家の宝刀」になっていると指摘している。

- 7 企業倒産や雇用調整による人員削減のうち、どの程度が不良債権処理の直接の影響によるものなのかは明らかでないために、不良債権処理が雇用に与える影響を示すデータは政府にもないが、内閣府の「不良債権の処理とその影響について」(2001年6月)によると、職を変えなければならなくなる人が39万人から60万人で、その内失業することになる人は13万人から19万人と推計されていた。
- 8 この緊急雇用対策交付金に対する評価を検討するが、対象とする年度を絞り、2002年度の評価を検討する。2001年度の評価は、緊急地域雇用対策交付金が2001年10月に創設されたものであるため支給実績が出ていないことに加え、2003年度以降の評価書によって、2002年度の評価書の見直しや改善についての変化が検討できるためである。
- 9 各府省は独自に評価計画を作っており、厚生労働省の評価の枠組みは、「厚生労働省政策評価実施要領」(2001)、「政策評価運営方針」(2001)、「厚生労働省における政策評価に関する基本計画」(2004)を参考にした。
- 10 厚生労働省は、評価の客観性担保のために、第三者機関である「政策評価に関する有識者会議」(以下、有識者会議)を設置し、審議しているが、個別の事業に対する評価の客観性についての議事はなされていない。この問題の評価が行われた2003年当時の議事内容は、制度の枠組み検討など厚生労働省の評価制度全般についての適切さなどを議論が中心であった。
- 11 総務省行政評価局『政策評価の点検結果』(2006年3月)では、認定関連活動の報告として23の事例が報告されている。
- 12 この点は、2003年11月13日付朝日新聞東京版19面で指摘されている。
- 13 2003年11月26日に行なわれた経済財政諮問会議に臨時委員として出席した坂口厚生労働大臣は、「一般に雇用対策については、アウトカム目標を明確にする必要があるのではないか」との考えを述べた。
- 14 これらの事実に対しては2003年4月22日の参議院厚生労働委員会での山本たかし議員が質問を行なった。そこでは、緊急雇用助成金を受けた事業所で継続雇用に結びついている人数(23,000人のうち362人)を確認したうえで、雇用創出への効果を検証する必要性が主張された。
- 15 この指摘は2003年11月13日付朝日新聞東京版1面を参照した。
- 16 この指摘は2003年10月13日付朝日新聞東京版2面を参照した。
- 17 2004年11月9日付け会計検査院決算報告においても「緊急雇用創出特別基金事業」と「早期再就職支援基金事業」の不適正支給の実態が指摘され、労働局による調査が不十分であったとも指摘している。
- 18 この指摘は2003年12月2日付朝日新聞東京版22面を参照した。
- 19 大阪府緊急雇用創出事業のHP (<http://www.pref.osaka.jp/koyosuishin/kikin/index.html>)(2006年7月)を参考にした。
- 20 以下の総務省の照会の内容と厚生労働省の回答に関する記述は、総務省資料「緊急地域雇用創出特別交付金(実績評価)について」を基にしている。
- 21 つまり、朝日新聞の調査によると厚生労働省が継続して雇用されたと主張していた13%のうち、事業所で常用雇用されたものは1.6%であり、残りの11.4%はパート・アルバイトによる雇用であったことになる。
- 22 大臣発言にまで至った雇用対策の施策目標に対する数値目標の設定に関する事項は、2004年度厚生労働省の政策評価書を見ても改善がなされていない。

参考文献

- 外務省経済協力局評価室(2003)「ODA評価ガイドライン」、外務省経済協力局評価室
- 厚生労働省(2002)「厚生労働省における政策評価に関する基本計画」、厚生労働省
- 総務省行政評価局(2003a)『評価の実施の必要性の認定』の考え方の整理と今後の取り組み」、総務省行政評価局
- 総務省行政評価局(2003b)「各府省が実施した政策評価についての審査の総括報告 14年度総括」、総務省行政評価局
- 総務省行政評価局(2005)『政策評価見直しの論点整理』、総務省行政評価局
- 総務省行政評価局(2006)『政策評価の点検結果 評価の実効性の向上に向けて』、総務省行政評価局
- 田辺国昭(2001)「政策評価制度の構築とその課題」、『日本労働研究雑誌12月号』、独立行政法人労働政策

研究・研修機構
南島和久、(2004)「『政策』と『評価』の間」、2004年
度日本公共政策学会報告論文
平松英哉(2005)「政策評価制度の機能不全 設計ミス
が生んだ悲劇」『社会科学』同志社人文学研究所
古川俊一(2005)「政策評価にみる本質的矛盾 - アカ
ウンタビリティと経営」、日本評価学会第6回全国大
会発表要旨

山谷清志(2000)「評価の多様性と市民」、『行政の評
価方式に関する調査研究』、行政管理研究センター
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』、第2巻
第2号、日本評価学会
山谷清志(2006)『政策評価の実践とその課題 アカ
ウンタビリティのジレンマ』、萌書房
(2007.1.30受理)

A Study on Activities to Ensure Objectivity of Evaluation in Japan

Tetsuya Nishimoto

Graduate School of Policy and Management, Doshisha University
ebg1012@mail3.doshisha.ac.jp

Abstract

In 2001, the Japanese government tried to introduce a 'safety-net' for the unemployed, especially the Special Emergency Grants for Job Creation. But there have been doubts whether these grants projects were effective or not. So, the press and the Diet queried to the projects, but the Ministry of Health, Labour and Welfare, which had responsibility for the projects, insisted the project performed well.

The function of meta-evaluation was adopted in the evaluation system of the Japanese government, and the Administrative Evaluation Bureau plays the main role objectivity of evaluation. In the Special Emergency Grants for Job Creation, objectivity of evaluation system became a problem and resulted in disputes.

Keywords

meta-evaluation, objectivity of the evaluation,
the administrative evaluation bureau, Special Emergency Grants for Job Creation

【研究ノート】

Trends in Air Pollutant Concentrations and Related Environmental Legislation in India

Miyuki Nagashima

Osaka School of International Public Policy, Osaka University, Osaka, Japan
mnagashi@osipp.osaka-u.ac.jp

Abstract

One of the major environmental problems facing India today is air pollution. The Indian government has already taken a number of measures to control air pollution. The purpose of this paper is to examine the impact of policy instruments on controlling air pollution in industrial areas in India from 1991 to 2002 with particular reference to trends in the concentration of SO₂, NO₂ and SPM. The results indicate that the concentration of SO₂ at all monitoring stations in 2002 met India's National Ambient Air Quality Standards (NAAQS) as a result of measures such as the diesel-sulfur phase-out program. Also, with regard to the concentration of NO₂, the fraction of stations reporting levels above the NAAQS has gradually declined since 1997, and the national annual average of NO₂ has been within the NAAQS at the majority of monitoring stations. SPM concentrations, however, have shown no such decline. One reason for this may be that the coal produced in India has a high ash content. This paper implies that because the SPM emission levels depend on the ash content of coal, stricter measures for controlling SPM will be needed.

Keywords

air pollution, environmental legislation

1. Introduction

One of the major environmental problems identified in India is air pollution. Delhi is one of the 10 most polluted cities in the world. Epidemiological studies indicate that high levels of air pollutants such as sulfur dioxide (SO₂), suspended particulate matter (SPM), and nitrogen dioxide (NO₂) have adverse effects of health, such as breathing problems and death

from respiratory and cardiovascular conditions. Cropper et al. (1997) estimated the relationship between air pollution and mortality in Delhi, the capital of India, and concluded that air pollution is a statistically significant determinant of daily deaths, with the exception of deaths in the age groups 0-4 years and 65 + years (Cropper et al. 1997).

Due to the rapid industrialization of India, sources of air pollution include industrial plants, motor

vehicles, refuse burning, and households burning biomass for cooking and heating. Recently, it has been reported that vehicles are major contributors to air pollution in urban areas, and high concentrations of industry and thermal power plants are also major causes of air pollution in other areas (United Nations Environment Program 2001). Transport has become a major source of urban air pollution, and recently the number of motor vehicles in particular has been increasing, resulting in a rise in vehicle emissions. According to the Asian Development Bank, the number of cars in India has doubled every seven years for the past 30 years (Asian Development Bank 1999). Such an increase in the number of vehicles has led to a higher level of energy consumption, traffic congestion and air pollution level. In the capital city, as much as 70 % of the air pollution is attributed to vehicular emissions (Central Pollution Control Board 1999). In addition to the increase in the number of cars, poor roads, fuel quality and vehicle maintenance also exacerbate air pollution in urban cities.

With regard to SPM, the major causes of emissions are the increase in the number of vehicles, the extensive use of fossil fuel in industrial activities and thermal power, wind-blown dust from agricultural processes and dust from unpaved roads. The sulfur content of diesel has a direct impact on the SO₂ and SPM emissions as shown by, for example, an Auto / Oil study implies that the relationship between particulates and sulfur level was found to be linear and that lowering the sulfur level in diesel would lead directly to a reduction in particulates from vehicles (Central Pollution Control Board 1999). The government has therefore taken some measures that tighten auto-emission norms (1991, 1996, 1998 & 2000), and has recommended fuel-quality specifications (1996) through the Central Pollution Control Board (CPCB). CPCB has also introduced a series of Action Points (1995) and a Strategy for Vehicular Pollution Control (1997) to control vehicle exhaust. Air pollution control legislation in India will be discussed further in the following section.

The purpose of this paper is to examine the

effectiveness of policy instruments in controlling air pollution in industrial areas in India from 1991 to 2002 by looking at the concentration trends of SO₂, NO₂ and SPM. At the time of writing, data concerning air pollutants in industrial, residential and sensitive areas are available from 1987 to 2003 on the CPCB website (Central Pollution Control Board 2004). However, due to the small number of samples from 1987 to 1990 and 2003, we will restrict our analysis to the trends in these concentrations from 1991 to 2002. Furthermore, we will limit our study to an examination of air pollutants in industrial areas. This is because more serious problems of air quality may be expected in industrial areas, where the air quality standards are less strict than in other areas, and the impact of economic activities on air quality would seem to be highly reflected in the industrial areas.

This analysis will assess the effectiveness of the measures taken to prevent air pollution in India. As far as I could determine, there are few papers or reports that analyze the trends of major air pollutants over the past decade and use data from observation points throughout India to examine what percentage of the observation points in India were over the National Ambient Air Quality (NAAQ) and WHO standards. The trends found in this analysis will shed light on the effectiveness of environmental legislation in India in improving air pollution. The paper is organized as follows: Section 2 describes the legislation and regulation related to air pollution control; in Section 3, the economic trends in India will be discussed; Section 4 provides data sources and examines the trends in the concentrations of air pollutants; Section 5 concludes.

2. Environmental Legislation in India

2.1 Central Pollution Control Board and Air Quality Standards

In September, 1974, the CPCB was established as a legal institution under the Water (Prevention and Control of Pollution) Act. Later, the Air (Prevention and Control of Pollution) Act of 1981 gave the CPCB

the function of preventing and controlling air pollution and the power to fulfill this function. According to the Air Act, one of the main functions of the CPCB with respect to air pollution is to lay down standards for emissions from automobiles, industrial plants, and any other source of pollution in order to improve the air quality and control air pollution in India. In 1984, the CPCB established a nation-wide program, called the National Ambient Air Quality Monitoring (NAAQM) Program, with 7 monitoring stations at Agra and Anpara. By 2000-2001, the number of monitoring stations had increased to 295, covering 99 cities / towns in 28 states and 4 Union Territories in India. Under this program, four types of pollutants, such as sulfur dioxide (SO₂), nitrogen dioxide (NO₂), suspended particulate matter (SPM), and respirable particulate matter (RPM (size less than 10 microns)) are regularly monitored at all of the stations, and National Ambient Air Quality Standards (NAAQS) have been set down for industrial, residential /rural/other areas, and sensitive areas with respect to major pollutants, as shown in Table 1. The CPCB has continued to provide daily and annual data on air

pollution collected at the monitoring stations in India.

Table 2 shows the WHO Air Quality Guidelines for SO₂ and NO_x, which provide guidance for countries in setting their national air quality standards. The guidelines define a concentration of the air pollutant below which no adverse effect on human health is expected.

The available evidence and information do not include guideline values for the concentration of particulate matter. For this reason, no guideline values have been recommended. For the purposes of this paper, we will take the National Ambient Air Quality Standards in the U.S. for annual average of PM₁₀ (particulate matter smaller than 10 micrometers) (50 $\mu\text{g}/\text{m}^3$) as a world standard.

2.2 Measures for vehicular exhaust

The government has implemented a number of measures to control emissions from vehicles. The major policy instruments intended to curb vehicular pollution and their current status are described in Kathuria (Kathuria 2002). Some of the instruments relate to engine modifications or technology, while

Table 1 National Ambient Air Quality Standards (NAAQS): Annual average ($\mu\text{g}/\text{m}^3$)

Pollutant	Industrial areas	Residential / rural and other areas	Sensitive areas
Sulfur Dioxide (SO ₂)	80	60	15
Oxides of Nitrogen as (NO _x)	80	60	15
Suspended Particulate Matter (SPM)	360	140	70

Source: Central Pollution Control Board 1999

Table 2 WHO Guidelines for SO₂ and NO₂

Pollutant	Time-weighted average ($\mu\text{g}/\text{m}^3$)	Averaging time
SO ₂	125	24 hours
	50	1 year
NO ₂	200	1 hour
	40	1 year
Particulate matter	-	-

Source: World Health Organization Regional Office for Europe 2000

Table 3 Diesel sulfur phase-out program in India

Phase	Target date	Sulfur content	Target area
Phase	April 1996	Low Sulfur (0.5%)	Four metros and Taj Trapezium
Phase	August 1997	Low Sulfur (0.25%)	Delhi and Taj Trapezium
Phase	April 1998	Low Sulfur (0.25 %)	Metro Cities
Phase	April 1999	Low Sulfur (0.25 %)	Entire Country

Source: Central Pollution Control Board 1999

other measures tighten emissions norms by applying European emission standards: EURO ² for heavy diesel vehicles and EURO ³ norms for diesel passenger cars. The fact that the sulfur level in diesel fuel is closely related to the emission of SO₂ and particulates was the main driver behind the measure to reduce the sulfur content in diesel fuel. As shown in Table 3, the sulfur content in the diesel fuel supplied to four metro areas and Taj Trapezium was reduced to 0.5% in 1996 (Phase 1), and was further reduced to 0.25 % in 1997 in Delhi and Taj Trapezium (Phase 2). The sulfur content in the diesel fuel (0.25%) was, according to government plans, to be available throughout India by April 1999. Finally, a measure requiring the desulfurization of petrol was implemented with a target date of May 31, 2000.

The regulation of sulfur levels is stricter internationally than in India. For example, some European countries have already introduced or have been preparing for sulfur-free (less than 0.001% of sulfur) diesel fuel. Further restrictions on the sulfur content in diesel fuel will be necessary to control air pollutants in India.

3. Economic trends in India

This section discusses economic trends in India from 1991 to 2002, before examining trends in air pollutant concentrations, since economic activities have a significant impact on air quality.

Figure 1 shows India's GDP growth rate expressed as percent change over the previous year,

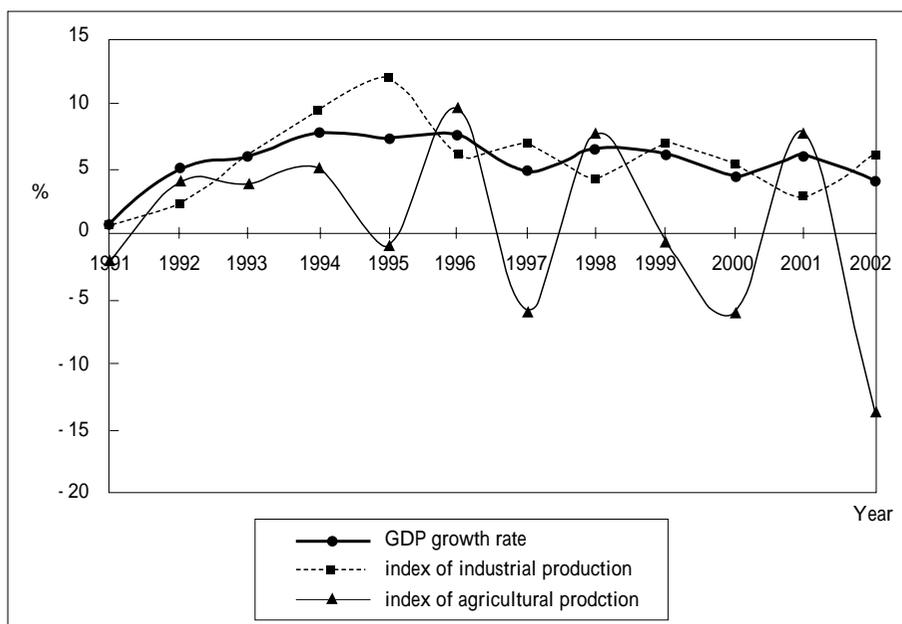
and presents an index of both industrial and agricultural production. According to the Economic survey 1991-1992, GDP growth decelerated significantly in the year 1991-1992, giving a final rate of 0.8 % growth. This was mainly due to the balance of payments problems which began in 1990-1991. To tackle this economic crisis, the government took measures to stabilize the economy and implement structural reforms.

Subsequently, as a result of these reforms, the economic situation gradually improved along with the growth in industrial production until the year 1996-1997. In the year of 1997-1998, the GDP growth rate decreased due to the sharp decrease in the agricultural production. The economic recovery in 1998-1999 was led mainly by a sharp increase in agricultural production. However, industrial production decelerated, which may have been the result of the East Asian crisis in 1997 and economic sanctions by the U.S. following Indian nuclear tests in 1998. Since 1999, agricultural production has fluctuated and industrial production decreased before increasing again in the year 2002.

4. Overview of Trends of Air Quality in India

This section presents comprehensive data on the concentrations of SO₂, NO₂ and SPM and their sources, and describes summary statistics regarding air pollutants in India.

Figure 1 GDP growth rate, index of industrial production and agricultural production in India from 1991 to 2002



Source: Government of India, Ministry of Finance, Economic survey 1994-1995, 1995-1996, 1996-1997, 1997-1998, 1998-1999, 1999-2000, 2000-2001, 2001-2002, 2002-2003, 2003-2004 and 2004-2005

Note:

1. Data in 2002 are provisional.
2. GDP figures are at 1980-1981 prices from 1991 to 1994 and at 1993-1994 prices from 1995 to 2002.
3. Index of agricultural production with base triennium ending 1981-1982=100.
4. Index of industrial production 1980-1981=100 in the data from 1991 to 1994. Index of industrial production 1993-1994=100 in the data from 1995 to 2002.

4.1 Data

Yearly data for the concentrations of SO₂, NO₂, and SPM are available on the CPCB website (Central Pollution Control Board 2004). Due to the small number of samples from 1987 to 1990, our analysis runs from 1991 to 2002.

Table 4 shows the descriptive statistics of the data on air pollutants in industrial areas from 1991 to 2002. The coefficient of variation, which is defined as a ratio of the standard deviation to the mean, for SPM is relatively smaller than that for SO₂ and NO₂. This implies that the dispersion of the SPM values is less

Table 4 Descriptive statistics for the data on the main pollutants in industrial areas from 1991 to 2002

Pollutant	Observation	Mean ($\mu\text{g}/\text{m}^3$)	Standard deviation ($\mu\text{g}/\text{m}^3$)	Coefficient of variation	Min. ($\mu\text{g}/\text{m}^3$)	Max. ($\mu\text{g}/\text{m}^3$)
SO ₂	769	22.7	21.7	0.96	0	191
NO ₂	775	32.0	29.5	0.92	0	273
SPM	736	233.9	133.75	0.57	16	876

Source: Own calculation based on data from Central Pollution Control Board (Accessed in 2004).

than other air pollutants. In the following section, we examine the change in the concentration of each pollutant in detail.

It should be noted, first of all, that the total number of monitoring stations with data differs from year to year due to missing data from between 1991 and 2002. Secondly, few dramatic compositional changes in the data from 1991 to 2002 have occurred. In other words, monitors in cities which became cleaner have not been removed and have continued to observe during the study period. This is not the case in the U.S., where the Environmental Protection Agency tends to remove monitoring stations from air “ clean ” counties and to add monitors in clean “ dirty ” counties (Greenstone 2004).

4.2 Trends in air pollutant concentrations

4.2.1 Sulfur dioxide

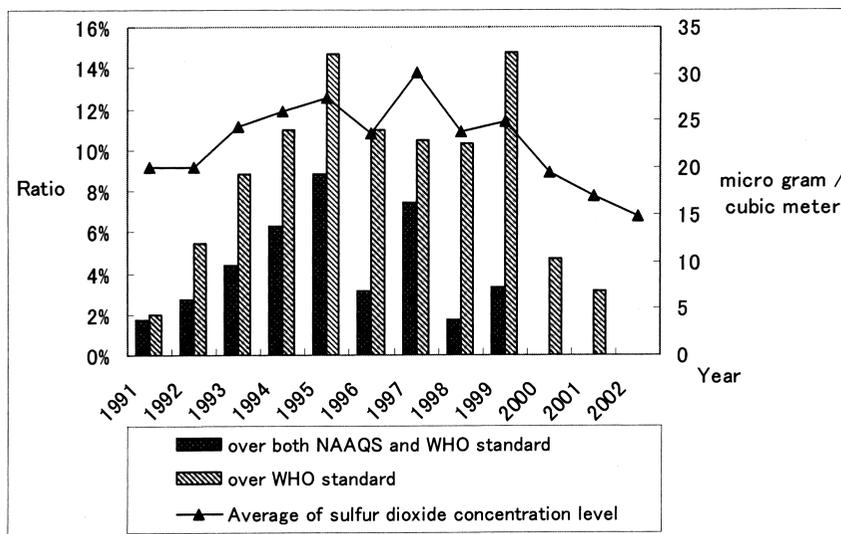
Figure 2 implies that the national average SO₂ concentration (shown as a line plot in Figure 2) across all monitoring stations meets both standards from 1991 to 2002. When we look at the fraction of stations

which violate the NAAQS (80 μg/m³), approximately 9 % of monitoring stations violated the NAAQS in 1995. In 2002, the concentrations of SO₂ at all monitoring stations met NAAQS and WHO standard.

The average SO₂ concentration achieved a peak of 30 μg/m³ in 1997, and then declined gradually to a level of 15 μg/m³ in 2002. Indeed, the average SO₂ concentration level has decreased since 1999. This finding suggests that the measures adopted to enforce cleaner fuel quality, such as changing to lower sulfur content in fuel (0.25%) in the entire country by April 1999, resulted in the lower SO₂ level. In addition, the requirement for the desulfurization of petrol (0.05%) with a target date of May 31, 2000, may have led to the drop in the concentration of SO₂ to within the NAAQS (80 μg/m³) across all monitoring stations after 2000.

Furthermore, the concentrations have decreased significantly since 1998. This could be the result of a fall in industrial production due to the East Asian crisis and the economic sanctions by the U.S., as discussed in Section 3⁴.

Figure 2 Histogram showing trends in the average annual mean SO₂ concentrations across all monitors in industrial areas (1991-2002)



Source: Own calculation based on data from Central Pollution Control Board (Accessed in 2004).
 Note: The WHO standard used above is 50 μg/m³ and the National Ambient Air Quality Standard (NAAQS) is 80 μg/m³.

4.2.2 Nitrogen dioxide

Figure 3 reports the trends in average annual NO₂ concentrations in industrial areas. Although about 8.8 % of the monitoring stations exceeded the NAAQS in 1995, the fraction of the stations over the standard has gradually declined since 1997, and the national annual average of NO₂ has been within NAAQS levels at most monitoring stations. The concentrations decreased overall in 1998 due to the fall in industrial production, as with the figures for SO₂. However, the fraction appears to increase after 2002, with 4.3 % of the stations over the NAAQS.

4.2.3 Suspended particulate matter

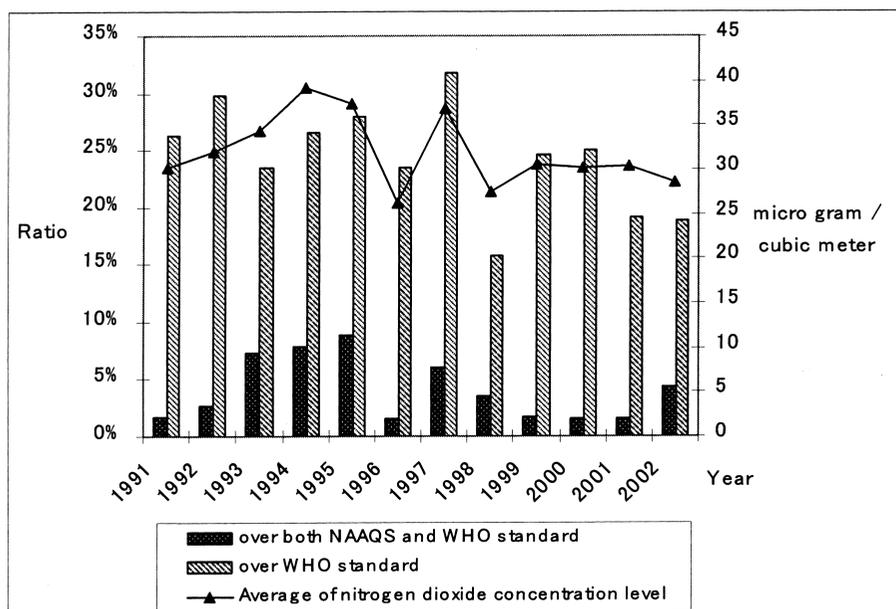
Figure 4 shows that the percentage of monitoring stations exceeding the NAAQS reached 20 % in 2002.

Although the government has formulated a variety of measures including legislation, guidelines, and air quality standards for air pollution and vehicular pollution, the concentration of SPM remains a major

problem. The analysis shows that the policy instruments used to prevent air pollution have not yet resulted in an improvement in SPM concentrations.

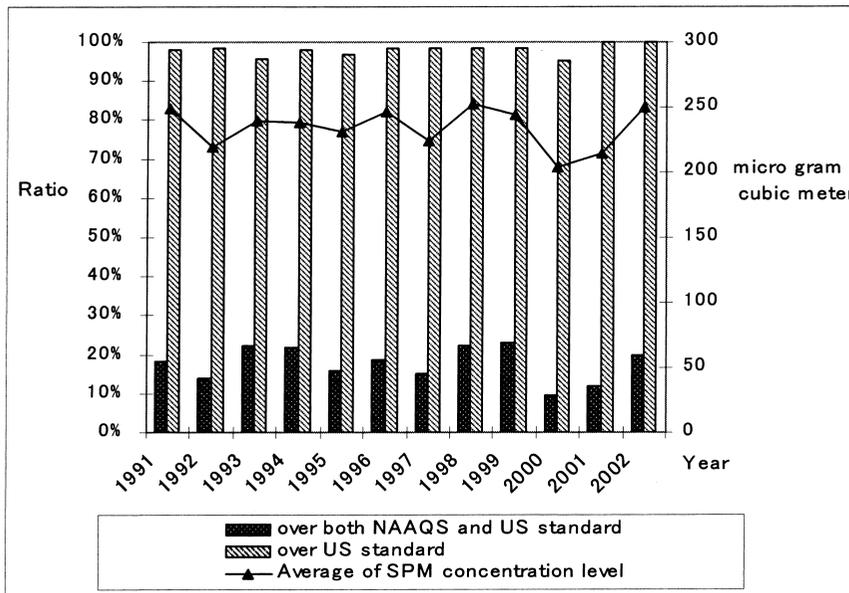
One possible reason that SPM concentrations have not declined might be the content of the coal produced in India. Chaulya et al. (2001) showed that haul routes and ore processing plant were the main sources of SPM emissions. According to the Ministry of Coal & Mines, Department of Coal, in India, coal is the primary fuel for energy in India, accounting for 55 % of India's energy use. India is the third largest producer of coal in the world, and it produced 337 million tons during the years 2002-2003 (Ministry of Coal 2005). Coal has been the primary source of energy for power production in India, and about 70% of coal is consumed by power generation (Central Pollution Control Board 2000). The thermal power sector, largely based on coal, makes a particular contribution to air pollution in India. Due to the impact of the diesel sulfur phase-out program in India,

Figure 3 Histogram showing trends in the average annual mean NO₂ concentrations across all monitors in industrial areas (1991-2002)



Source: Own calculation based on data from Central Pollution Control Board (Accessed in 2004).
 Note: The WHO standard used above is 40 µg/m³ and the National Ambient Air Quality Standard (NAAQS) is 80 µg/m³.

Figure 4 Histogram showing trends in the average annual mean SPM concentrations across all monitors in industrial areas (1991-2002)



Source: Own calculation based on data from Central Pollution Control Board (Accessed in 2004).
 Note: The US standard used above is 50 µg/m³ and the National Ambient Air Quality Standard (NAAQS) is 360 µg/m³.

Indian coal generally has low sulfur content: less than 0.4%, but a high ash content ranging from 24% to 45% (Central Pollution Control Board 2000).

Under the notification of the Ministry of Environment and Forests of the Government of India on September 14, 1999, the utilization of ash from thermal power plants was mandated beginning in June, 2000. As the data used in this paper covers the period up to and including 2002, it can be said that this mandate had not yet taken full effect. Since the quantity of SPM emissions depends on the ash content of the coal, stricter measures for controlling SPM may be needed.

5. Conclusion

Some Indian cities have a serious air pollution problem. The government has implemented a number of environmental instruments to control air pollution. Owing to the measures controlling sulfur emissions,

such as the diesel sulfur phase-out program and the instrument for the desulfurization of petrol, SO₂ levels at all monitoring stations met NAAQS in 2002. However, although the national average SPM concentration level in 2002 is under the NAAQS, and amounted to about 250 µg/m³, about 20% of the monitoring stations were still over the NAAQS. One of the reasons for this result may be that the coal produced in India includes a high ash content, ranging from 24 to 45% (the quantity of SPM emissions depends on the ash content of the coal). In addition to the sulfur phase-out programs and the ongoing utilization of ash from thermal power plants, further strict measures for controlling SPM may be needed.

Acknowledgement

The author would like to thank Kanemi Ban and an anonymous referee for constructive comments on this paper.

Notes

- 1 The annual standard for PM₁₀ was eliminated, which was effective in December, 2006. For more information, refer to Esworthy R. and McCarthy J.E. (2006).
- 2 It limits emissions of NO_x to 0.5 g/km and of Particulate Matter (PM) to 0.05 g/km for diesel car emissions and emissions of NO_x to 0.15 g/km for gasoline car emissions (see European Union(1998)).
- 3 It limits emissions of NO_x to 0.25 g/km and of Particulate Matter (PM) to 0.025 g/km for diesel car emissions and emissions of NO_x to 0.08 g/km for gasoline car emissions (see European Union (1998)).
- 4 We got valuable comments from the anonymous referee in this context.

References

- Asian Development Bank (1999). *Urban Sector Strategy*. Manila: Asian Development Bank.
- Central Pollution Control Board (1999). Auto Emissions. *Parivesh Newsletter June*. New Delhi: Central Pollution Control Board.
- Central Pollution Control Board (2000). Clean coal initiatives. *Parivesh Newsletter June*. New Delhi: Central Pollution Control Board.
- Central Pollution Control Board (Accessed in 2004). Ambient Air Quality in India. (<http://www.cpcb.nic.in/data.htm>)
- Chaulya, S.K., Chakraborty, M.K. and Singh, R.S. (2001). Air pollution modelling for a proposed limestone quarry. *Water, Air, and Soil Pollution*, 126. 171-191.
- Cropper, M.L., Simon, N.B., Alberini, A., Arora, S., and Sharma, P.K. (1997). The health benefits of air pollution control in Delhi. *American Journal of Agricultural Economics*, 79 (5), 1625-1629.
- Esworthy, Robert and McCarthy, James E. (2006). *Air Quality: EPA's 2006 changes to the Particulate Matter (PM) standard*. CRS Report for Congress, Congressional Research Service, The library of Congress. Washington D.C. (electronic resource).
- European Union (1998). *Official Journal of the European Communities* (OJL 350). European Union.
- Greenstone, Michael (2004). Did the Clean Air Act cause the remarkable decline in sulfur dioxide concentrations?. *Journal of Environmental Economics and Management*, 47, 585-611.
- Kathuria, Vinish (2002). Vehicular pollution control in Delhi. *Transportation Research Part D*, 7 (5), 373-387.
- Ministry of Coal, Government of India (2005). *Induction Material*. Chapter3. Ministry of Coal.
- Ministry of Finance, Government of India (1992). *Economic survey 1991-1992*. New Delhi: Ministry of Finance.
- Ministry of Finance, Government of India (1995). *Economic survey 1994-1995*. New Delhi: Ministry of Finance.
- Ministry of Finance, Government of India (1996). *Economic survey 1995-1996*. New Delhi: Ministry of Finance.
- Ministry of Finance, Government of India (1997). *Economic survey 1996-1997*. New Delhi: Ministry of Finance.
- Ministry of Finance, Government of India (1998). *Economic survey 1997-1998*. New Delhi: Ministry of Finance.
- Ministry of Finance, Government of India (1999). *Economic survey 1998-1999*. New Delhi: Ministry of Finance.
- Ministry of Finance, Government of India (2000). *Economic survey 1999-2000*. New Delhi: Ministry of Finance.
- Ministry of Finance, Government of India (2001). *Economic survey 2000-2001*. New Delhi: Ministry of Finance.
- Ministry of Finance, Government of India (2002). *Economic survey 2001-2002*. New Delhi: Ministry of Finance.
- Ministry of Finance, Government of India (2003). *Economic survey 2002-2003*. New Delhi: Ministry of Finance.
- Ministry of Finance, Government of India (2004). *Economic survey 2003-2004*. New Delhi: Ministry of Finance.
- Ministry of Finance, Government of India (2005). *Economic survey 2004-2005*. New Delhi: Ministry of Finance.
- United Nations Environment Programme (2001). *State of Environment, India 2001*. United Nations Environment Programme.
- World Health Organization Regional Office for Europe (2000). *Air quality guidelines for Europe second edition*. WHO Regional Publications European Series, No. 9.

【研究ノート】

多様化する学士課程教育の自己評価方法

串本 剛

広島大学大学院教育学研究科・日本学術振興会特別研究員

t.kushimoto@hiroshima-u.ac.jp

要 約

本研究の目的は、評価方法に関する4つの類型を設定し、学士課程教育の自己評価におけるそれらの採否状況について、全国調査から得られたデータを用いて実証的に明らかにすることである。前半部分では、学士課程に関する教育観の違いを反映するものとして、自律型、応需型、標準化型、成果準拠型という4つの評価方法を導入する。後半では、全国1,871の学部を対象として実施した質問紙調査の結果を使って、各評価方法の採用状況を実証する。評価方法の採用率は、評価領域や学部の属性により異なることが明らかになると共に、その違いが評価方法の背後にある教育観との関係から考察される。

キーワード

学士課程教育、自己評価、評価方法、多様化、全国調査

1. 研究の背景と目的

(1) 研究の背景

制度的かつ恒常的な取り組みとしての評価は、近年、産官学を問わず広く関心の対象となってきた（牟田 2001）。「学」の一部をなす高等教育も例外ではなく、最近の動向で言えば、すべての大学が7年以内ごとに文部科学省の認証を受けた評価機関による評価を受けなくてはならないとする認証評価制度が、2004年から導入された。

現在、学士課程教育を対象とする認証評価機関には、大学基準協会、大学評価・学位授与機構、日本高等教育評価機構の3つがあり、評価基準や項目、費用等の面では異なる枠組みを利用してい

る¹。ただし共通点もあり、そのひとつが大学の実施する自己評価に基づいて評価を行うという点である。自己評価を基礎とする第三者評価という仕組みは、日本に限らず世界的にも見られるものであり、それだけ大学にとって組織の自律性が重要であることを示していると言える。特に日本では、自己評価の制度化（1991年）の後、その一定の普及を経てから外部評価の努力義務化（1999年） 認証評価制度の導入と展開してきており²、今日でも大学をめぐる評価において自己評価が中心に位置づけられていると見ることができる。

自己評価とは、評価のプロセスにおける価値判断の主体がプログラムの実施者である、ということの意味するに過ぎないが、学士課程教育を対象としたそれは従来、価値判断だけ無くその際に考

慮される情報（「根拠情報」）すらも、主として評価対象に対する「自己」の主観的判断に負っていた³。例えば教育内容について評価しようとするとき、それについて自分たちがどう思うかということが、主に考慮されてきたのである。

しかし、各種の大学改革や上記の第三者評価の進行、あるいは大学を取り巻く社会経済的環境の変化に伴い、いまでは授業評価等を通して学生の意見が顧慮されるだけでなく、外部評価によって学外の主体による判断を参考にする機会も増えている。さらに教育の目的が達成されているかという根本的な問いに回帰し、教育成果に関する客観的・主観的な情報を把握し、それらに基づいて教育の是非を判断するという動きも見られる。つまり、評価方法の多様化が起きているのである。

（2）本論の目的と構成

こうした状況の推移が、「何を持って教育の是非を判断すべきか」という教育観の多様化とも並行するものであるとすれば、自己評価の実態解明には、単なる現状把握以上の意義があると考えられる。だが第三者評価のそれに比べ、大学の自己評価に関する研究蓄積は限られており、根拠情報との関連からその評価方法を明らかにするという試みは、世界的に見てもこれまでのところ為されていない⁴。

そこで本稿では、幾つかの異なる教育観を反映するものとして評価方法の類型を設定し、学士課程教育の自己評価におけるそれらの採用状況について、全国調査から得られたデータを用いて実証的に明らかにすることを目的とする。この結果得られる知見には、これまでブラック・ボックスであった現象に光を当てることだけでなく、変革期にある学士課程教育に関する議論にも、ひとつの視点を提供することで一定の貢献をすることが期待できる。

以下、第2節ではまず、分析の視点となる4つの評価方法の相互関係と、それぞれが依拠する教育観、および評価方法採用の徴拠となる根拠情報について解説する。続く第3節では、2006年に実施した全国調査の結果を使って、評価方法の採否状況について検討する。分析に際しては、評価領域

による相違と共に、学部の属性と採否の関係についても検証する。最後に第4節では、学士課程に関する教育観の視点から分析結果を考察し、今後の課題を提示する。

2. 分析の視点～4つの評価方法

（1）4類型の関係

評価の類型化には、ふたつの方法がある。ひとつは、「評価とは何か」という評価の定義の違いをも反映させる形で行われるもので、この場合、各類型はある程度背反的であることが前提とされており、「モデル」と称されることが多い。（プログラム）評価論一般の研究では、こうした類型化がたびたび試みられており、近年のものとしては、22のモデル（アプローチ）を4つのカテゴリーにまとめて論じたStufflebeam（2001）や、哲学的（客観主義 - 主観主義）方法論的（量的 - 質的）、実践的（情報提供のみか価値判断を含むか、評価者が指導的な役割を果たすか援助的か、評価者の経験、評価対象に関する専門知識が必要か否か、依拠するモデルがひとつか状況に応じて選択的か）な相違を考慮した上でHouse（1983）を参考にしつつ、5つのアプローチを示したFitzpatrick et al.（2004）がある。

アメリカの先行研究には、高等教育を念頭に置いたものの中にも、プログラム評価論の知見を援用してこのような類型化を行っているものが見られる。Gardner（1977）は初期の代表的な研究で、評価の時系列的な進化に対応させて5つの評価モデルを提示している。また説明責任の遂行に加え資源配分の根拠として評価が利用されるようになった頃のConrad and Wilson（1985）では、Gardner（1977）を踏襲する形で、目標標準拠（objective-based）、応答的（responsive）、意思決定（decision-making）、鑑識眼（connoisseurship）という4つの評価モデルを挙げている。

これらに対し、高等教育に関する日本の先行研究では、何らかの評価の定義をしばしば暗黙の前提として、評価に関わる諸要素について類型化を試みる傾向にある。例えば関（1989）は、評

価の性格、時期、過程・対象、主体、機能、方法の各側面から、同様に大学基準協会(2000)では、評価の主体、対象、内容、時期、対象の動態、結果の効果という切り口から類型化を行っている。この方法の場合、類型間の排他性は前提とされておらず、ある評価が複数の類型に当てはまるということもあり得る。どちらの方法による類型化が有効かは、研究の目的や関心により異なり、カリキュラムに関する著作で評価モデルを論じている Stark and Lattuca (1997) のように、上記の2つの中間的な方法による類型化の例もある。

研究の背景として述べた通り、本稿の関心はあくまでも学士課程教育の自己評価にあり、その評価方法の多様性についても、完全な移行ではなく、伝統的なそれに対する追加的な導入を前提としている。従って、本論における類型化は、どちらかといえば後者の方法によるものであり、4つの評価方法の類型(アプローチ)は、基本的に並存するものとして提示される。

(2) 自律型

最初の評価方法は、学士課程教育が「所属教員の専門性や見識を尊重し、自律的に営まれるべきである」という教育観を前提とするもので、「自律型」と名づけることができる。この評価方法では、根拠情報として評価対象に対する所属教員の主観的な判断が用いられるため、場合によっては評価のための情報収集が組織的に行われず、ということもあり得る。

自律型アプローチは、最も伝統的な評価方法であり、自己評価が制度的に位置づけられる以前から採用されてきたものと言える。なぜならば、その背景には中世以来の大学の自治、そして19世紀にドイツで興った近代大学における学問の自由という理念が存在するからである。古くは王権や宗教への従属、近代以降で言えば国家による統制や産業界への過度な迎合は、それらの理念を脅かすものとされてきたため、大学には外界からの干渉を忌諱してきたという歴史的経緯がある。これが、学士課程教育の是非はその教員によって判断されるべきだという、教育観に結びついているものと考えられる。

こうした発想は、Eisner (2003) の言う鑑識と批評 (connoisseurship and criticism) に依拠しているものと考えられ、従来の類型化でもしばしば取り上げられてきた (Gardner 1977, Conrad and Wilson 1985)。通例Eisnerのモデルは、アメリカにおけるアクレディテーションのような外部評価を説明する概念として用いられてきたが、より日常的な現象に目を向ければ、多くの自己評価がその具体例だということがわかる。

また、常にプログラム全体として評価が行われるわけではなく、各教員の個別授業などにおける日常的な評価の積み重ねが、組織的な行為を代替し得るという側面に注目すると、Stark and Lattuca (1997) が挙げている「非公式の評価 (informal assessment)」に該当するとも言える。

自律型の評価方法は、学士課程教育という評価対象の専門性の高さを考慮した場合、有力なアプローチであることは間違いない。ただし Fitzpatrick et al. (2004) が指摘するように、非公式の評価は評価者の経験や臆見により偏向してしまう危険性がある。

さらに、専門性に依拠した評価自体の妥当性が問われる際には、正当化が難しいことも事実である。これは単に、もはや象牙の塔ではなくなったという社会的な位置づけの変移だけではなく、教員個人の省察は教育観の制約を受けるため改善につながるとは限らない (Trigwell and Prosser 1996) というような、学問的知見の広まりにも由来した意識転換の結果でもある。こうした価値観の動揺が、以下に挙げる3つのアプローチの台頭を促していると見られる。

(3) 応需型

応需型の評価方法は、「学生という「第一の顧客」の要求に応えるべきである」という教育観に根ざすもので、評価対象に対する学生の意見を根拠情報として用いる。

「第一の顧客」とは、非営利組織について論じたドラッカーとスターン (訳書 2000) の用語で、「ある活動によって生活が改善される人々 (p. 20) のことを指す。一般に何らかの行為 (事業プログラム) には、それによって益すると考えられる

「顧客」が複数考えられる。学士課程教育も例外ではなく、学生以外にも、その保護者や卒業生の雇用者、さらに活動を通して達成感を得るという意味では、教員や職員もそれに含まれる。だが、教育を行う一義的な意図が彼/彼女らの変容にあるという意味では、学生こそが「第一の顧客」と言える。

営利的な活動であれば、応需型アプローチは基本的な評価方法とさえ言えるものであろう。しかし大学においては、専門知識・技能を持たない者(学生)がその所有者(教員)からそれらを教授されるという関係性が支配的であったため、学生の意見が評価の根拠となるという発想はなじみにくかった。その名残は、授業評価が導入される時に繰り返される「学生に評価をする能力(資格)はない!」というお決まりの台詞からも窺うことができる。セイモア(訳書 2000)は、Garvin(1988)による質の定義を引きながら、こうした問題を鮮明に解説している。

しかし学生が権利の主体、それも教育というサービスを購入する消費者として認識されるようになるのに伴い、「第一の顧客」である学生の意見を考慮することの必要性が認められるようになってきている。学生が消費者として位置づけられるに至った経緯については、1980年代初頭にリースマン(訳書 1986)がアメリカの文脈において論じているが、最も大きな要因は大学が大衆化することにより、学生の関心や能力が多様になったことである。そうした傾向は今日では多くの国々で見られ、学生は権利と具体的な期待・要望を持っているのだから顧客である、という認識は広く認められている(Davis 2002)。実際にわが国でも、「学生を、大学にとって主たるクライアントであると位置づけ直し、そのニーズを汲み取り、そのうえで教育プランを設定する」(日本私立大学連盟 学生部会 2000、p. 203)ことの必要性が訴えられるようになってきている。

事実、「第一の顧客」の意見を根拠情報とすることに特徴を求める評価モデルは、比較的新しい研究にしか見られない。Stark and Lattuca(1997)では、これを学生中心評価(student-centered evaluation)と呼び、「様々な目標を持った学生の授業やプログラムでの経験を知ることにより、教

員と学生の目的意識の一致度を調べる」(p. 291)作業であると位置づけている。本研究で言う応需型アプローチは、必ずしもそのような評価目的を前提とするものではないが、学生の意見を考慮するという点では、学生中心評価に共通するものである。

(4) 標準化型

「機関を超えて、同等の教育を提供すべきである」という教育観が、第3の評価方法である「標準化型」を支持するものである。ここではプログラムの外部から得られる情報が考慮されるため、教育が標準的な形に収斂していく可能性が内包されている。

同等性や通用性に価値を置く教育観は、大学が学位という世界共通の資格を発行する唯一の機関であるとされることからわかるように、大学の根本的な理念のひとつであると言える。ただし、その担保の仕方は時代と共に変遷してきており、概括的に言えば、かつては大学あるいは課程の設置時に同等性を保証するための仕組みが設けられていたのに対し、近年では事後的にそれを行う傾向にあり、標準化型アプローチによる自己評価はその機能を担っているということが出来る。こうした移行は、高等教育セクターの肥大化の結果、行政の役割が変化することによって起こるものであり、設置基準の大綱化に伴う自己評価の努力義務化という事実によって、日本においては顕現している。

「評価対象の標準化に対する志向」という方法上の特徴は、従前の評価モデルには見られないもので、ある意味で学士課程教育の自己評価独特の方法と言える。しかし、具体的な根拠情報を考えると、例えば学外の専門家による見解を考慮する場合は、自律型アプローチのところ而言及したEisnerのモデルを応用しているものと捉えられる。このとき注意すべきは、学外の専門家にはその専門性もさることながら、広く事情に精通していることによる見解の普遍性が期待されていることである。換言すれば、彼らの見解を参照することにより、当該教育の通用性が保証されることになる。こうした力点の違いが、ひとつの評

価モデルの原理が、ふたつの評価方法の文脈で参照される原因となっている。

標準化型アプローチを支える教育観が、学士課程教育にとって本質的なものである反面、高等教育の大衆化によって、学士課程教育が多様な形態・内容を持たざるを得なくなる状況の中で、一元性を志向するその評価方法自体に限界があることも確かである。次項で取り上げる成果準拠型が必要となる背景には、そうした事情も存在すると考えられる。

(5) 成果準拠型

ここまで論じてきた3つの評価方法とは異なり、行為の是非をその結果によって判断しようとする「成果準拠型」アプローチは、「教育の妥当性はその成果によって判断されるべきである」という教育観から発生する。そのため根拠情報としては、何らかの指標によって把握された教育成果が求められる。

教育とは、教育者の意図を学習者のうちに実現することを目的とする行為であり、その意味で成果準拠型アプローチが前提とする教育観は、至って常識的なものであると言える。しかし応需型の背景のところでも論じたように、大学における教員と学生の関係は、必ずしも上記の意味での「教育的関係」ではなかったため、教育目的や目標については教育成果を明確にするということは、一般的ではなかった。

こうした状況が変わりつつあるのには、主としてふたつの原因が考えられる。ひとつは、1980年代以降世界的に見られる説明責任論の隆盛の中で、結果に対する関心が高まっているということである。そこでの焦点は、対外的に教育成果を示すことにあると考えられるが、それが結果として自己評価における教育成果の重視にも繋がっている。もうひとつ、構成主義的な学習理論が大学教育の文脈でも重視されるようになってきたため、教育そのものの質を見ることによりその成果を推認することの妥当性が問われている、ということも見逃せない。この点については、Barr and Tagg (1995) 等が大学教育のパラダイム転換として論じているほか、評価に関連してはHuba and Freed

(2000) やPalomba and Banta eds. (2001) が言及している。

成果準拠型アプローチは、Conrad and Wilson (1985) によると体系的に論じられているものとしては最も古い評価モデルである、目標準拠モデル (objective-based model) の萌芽的な形態と見ることが出来る。その祖がTyler (1949) に求められる目標準拠モデルでは、教育目標を操作的に定義することにより達成の測定を可能にし、目標の達成度をもって教育の質を判断するという構造が前提とされる。本論における成果準拠型アプローチは、ここまで体系的な構造を有するわけではないが、教育の評価に際しその成果を考慮する点において共通の特徴がある。

教育成果を根拠情報として用いることには、原理的な正当性を認められるだけでなく、教育そのものの多元性を正当化しようという利点もある⁵。だが、実用の難しさもしばしば指摘される所であり、その代表的なものは、教育成果が把握されただけでは何を改善すべきかがわからない、ということである。わが国でも教育成果への関心が高まる中、同様の批判が見られる (大山 2001)。しかしこの問題に関しては、プログラム・セオリーやロジック・モデルと呼ばれる、原因と結果の関係性を明確化することによる対応策も論じられてきており (例えばRogers et al. 2000) 教育を対象とした評価においても徐々に浸透してきている⁶。

3. 評価方法の採否

(1) 分析データ

前節で提示した4つの評価方法の採用状況を明らかにするために本論で用いるデータは、学士課程教育の自己評価の実態を解明することを目的として開発した「教育評価と教育改善に関する調査」の調査結果である。

調査票は、学部の特徴、教育評価体制等、教育評価の方法、教育評価の効果・影響、教育評価に関する回答者の意見、について聞く5つの部分から構成されている。教育評価の体制に関しては経年比較の可能性を考慮し広島大学・大学教育研究

センター編（1999）の調査を、効果・影響については適切な日本の先行研究がなかったため、アメリカで実施された類似調査であるInstitutional Support for Student Assessment Inventory（ISSA）を参考にした⁷。また本論の主題である評価方法に関する問は、串本（2006）でまとめている自己評価報告書の分析や、第三者評価機関の取組を基に設定した。

調査対象は2005年5月の時点で学士課程を持つ全国1,871の学部で、質問紙は4月21日を締め切りとして、2006年3月28日に発送された。4月末には督促はがきを送り、最終的に6月15日時点での回収率は31%であった（表1）⁸。

表1 調査の概要

	国立	公立	私立	全体
発送数	345	158	1,358	1,871
回収数	215	53	309	580
回収率	62.3%	33.5%	22.8%	31.0%

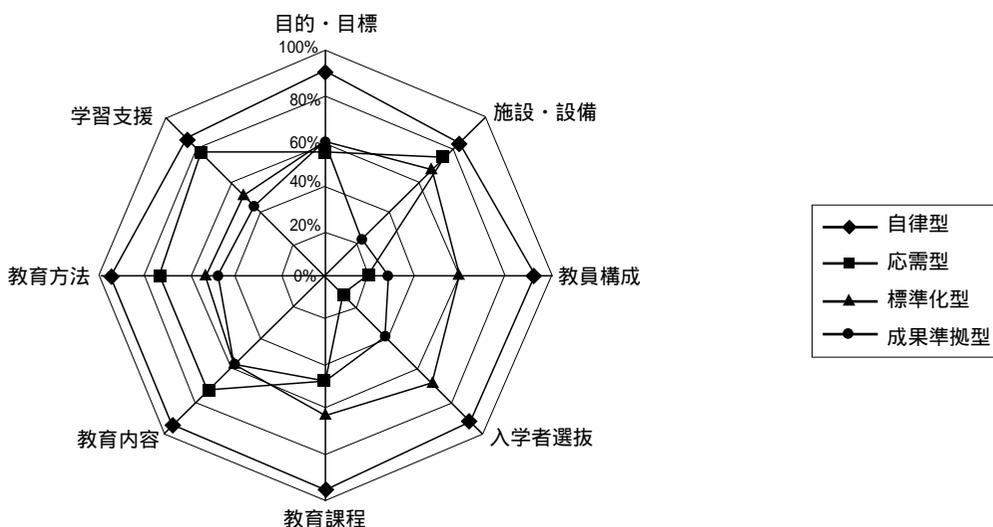
（出所）筆者作成

（2）全体的傾向

評価方法に関する問いでは、下記1）～8）の各領域について、現状の適否や改革の必要性を判断する際にどのような根拠情報を用いているかを、複数回答で尋ねた。その回答に基づき、各領域における評価方法の採用状況を示したものが図1である⁹。

- 1) 目的・目標：教育の理念・目的・目標の内容・具体性など
- 2) 施設・設備：教室、実験器具、情報インフラ、図書館など
- 3) 教員構成：職階・年齢・専門分野の構成、教育者としての素養など
- 4) 入学者選抜：入試の種類・内容など
- 5) 教育課程：コース・課程の設定、各種科目の配列・割合など
- 6) 教育内容：提供科目、授業内容など
- 7) 教育方法：授業方法、授業形態・規模、シラバス、成績評価方法など
- 8) 学習支援：履修指導（ガイダンス、オフィスアワー）、補習体制など

図1 評価方法の採用状況



（出所）筆者作成

自律型はいずれの領域でも8割以上の学部が採用しており、最も広く用いられている評価方法であることが確認できる。既述の通り評価方法は併用を前提としているので、むしろ自律型アプローチを「使わない」という事例は稀であると解釈した方がわかりやすい。

評価領域間で採用率に差がないという点では、標準化型の評価方法も共通している。採用率は全体に6割前後であり、普及が十分に進んでいるとは言い難いが、外部から獲得された情報は領域を問わず、評価の際に考慮されていることがわかる。

上記のふたつとは対照的に、応需型アプローチの採用状況には、評価領域によって格差がある。教員構成や入学者選抜の評価ではほとんど利用されていないのに対し、その他の6つの領域ではほぼ半数以上の学部が採用している。

成果準拠型の評価方法も、応需型に似て領域間で採用程度に差がある。また、採用率は全体に低く、最も普及している目的・目標の評価においても、それは6割に満たない。その成果を根拠として教育条件や教育活動の是非を判断するという方法は、大学においては依然として定着していないと言える。

全体的な傾向からは、自律型を基礎にしつつも評価方法の多様化が進んでいることが実証されたが、その様相は評価領域によってかなり異なることも明らかとなった。では、評価方法の採否には、学部の属性も関係するだろうか。次項でその点を論ずる。

(3) 評価方法の採否と学部の属性

評価方法の採否と関係があると考えられる学部属性として、ここでは設置形態、専門分野、卒業生の職業に関する想定の方針、入試難度の4つを設定した。専門分野は、学習における自由度の高さを反映すると考えられる文系・理系の別により、職業想定の方針は関連する問いの回答に基づいて、また入試難度は学部偏差値に従って下位と上位に区別した¹⁰。

表2には、評価領域毎に行った評価方法の採否と学部属性のクロス集計結果を、4つの評価方法についてまとめている。表中の値はすべて当該ア

プローチを採用している学部の割合で、*印はカイニ乗値による検定結果である。

まず自律型について見ると、いずれの領域でもほとんどの学部が採用していることもあり、有意差が見られる関係は多くない。目的・目標、教員構成、学習支援では設置形態、施設・設備においては専門分野や職業想定程度の違いで有意差が見られる。

応需型アプローチでは、有意差が見られるのは教育課程の評価についてのみだが、その他5つの領域でも理系より文系の方が採用率が高くなっている。専門分野以外を見ると、教育方法と学習支援は設置形態の違いで、施設・設備は職業とのつながりの強弱で採用率に開きが見られる。特に設置形態に関しては、教育課程を除くすべての領域で公立大学での採用率が低い。

標準化型の採否と学部属性の間には、一転して統計的に有意な関係が数多く確認できる。まず設置形態では、私立大学における採用率の低さが顕著であり、応需型の場合とは対照的に、公立大学で採用率が高くなっていることも特筆に値する。専門分野による違いも明らかで、ほぼすべての領域で理系学部の採用率が高くなっている。

最後の成果準拠型は、標準化型以上に職業との繋がりによる差異が大きく、設置形態や文理の別では有意差が見られない。これに加え、入学者選抜、教育内容、教育方法では入試難度による採用率の相違も見られる。

以上の分析からは、自律型や応需型に比べ、標準化型や成果準拠型では学部属性との間に顕著な連関が見られ、特に前者では設置形態や専門分野、後者では職業との繋がりという属性が、違いを鮮明に描き出したと指摘できる。

4. 考察と今後の課題

(1) 領域間格差について

ここまで示してきた分析結果に対し、各評価方法が前提としている教育観の視点から考察を加えると、幾つかの知見を提供することができる。

各評価方法の全般的な採用率からは、自律型を

表2 評価方法の採否と学部属性

(%)

	全体	設置形態			専門分野		職業想定		入試難度	
		国立	公立	私立	理系	文系	否定	肯定	下位	上位
自律型										
目的・目標	91.2	95.1	90.9	88.5*	92.0	91.1	94.0	90.4	92.9	90.1
施設・設備	84.0	88.2	84.8	80.9	89.1	81.6*	76.1	86.1*	82.5	86.1
教員構成	92.7	96.5	95.5	89.5*	93.2	92.2	89.2	93.6	92.4	92.8
入学者選抜	91.8	94.9	93.5	89.2	92.3	91.5	89.8	92.3	90.0	93.1
教育課程	95.5	96.5	100.0	94.0	94.4	96.2	97.4	95.2	95.5	95.1
教育内容	95.1	96.5	91.3	94.8	94.4	95.6	92.3	95.9	95.5	95.2
教育方法	95.0	95.5	91.3	95.1	94.9	94.9	94.0	95.2	95.5	94.0
学習支援	86.5	92.4	84.4	82.6**	86.2	86.3	86.4	86.7	83.8	88.6
心需型										
目的・目標	55.1	54.7	43.2	57.0	50.8	58.3	51.3	56.1	55.4	55.3
施設・設備	73.2	77.9	60.9	72.2	72.5	74.7	60.6	76.8**	73.1	74.4
教員構成	19.8	19.7	15.9	20.7	15.6	21.8	18.9	20.1	19.1	20.4
入学者選抜	12.3	14.6	10.9	10.8	12.4	12.4	16.7	11.1	11.7	13.4
教育課程	47.2	44.3	47.7	49.3	39.5	52.2**	47.4	47.3	50.2	45.7
教育内容	72.2	76.1	58.7	71.5	69.2	74.4	68.4	73.1	71.1	73.7
教育方法	73.8	80.6	67.4	70.0*	75.8	73.3	68.4	75.2	73.7	74.6
学習支援	77.7	84.8	71.1	73.7**	80.6	76.5	71.8	79.4	77.1	78.4
標準化型										
目的・目標	57.3	66.5	63.6	49.7**	65.8	52.9**	52.1	59.0	56.7	58.6
施設・設備	66.1	68.7	65.2	64.6	71.0	64.1	63.3	67.0	65.8	66.5
教員構成	59.0	67.2	70.5	51.1**	65.1	55.7*	51.4	60.9	56.4	61.9
入学者選抜	67.2	74.7	71.7	61.0***	67.0	68.3	63.0	68.3	65.0	70.6
教育課程	61.9	69.7	72.7	54.6**	67.7	58.6*	53.5	64.5*	56.7	67.2*
教育内容	55.8	63.2	69.6	48.3**	65.7	49.5***	45.3	59.0**	52.4	58.9
教育方法	52.5	56.7	54.3	49.5	62.6	47.6**	42.7	55.3*	47.8	58.6*
学習支援	51.0	53.8	60.0	47.7	53.1	50.7	42.7	53.4*	49.6	53.4
成果準拠型										
目的・目標	59.0	60.6	61.4	57.3	63.8	56.1	46.2	62.9**	57.9	60.8
施設・設備	22.7	24.6	21.7	21.7	25.9	21.1	13.8	25.2*	24.4	21.4
教員構成	27.9	27.8	25.0	28.6	27.6	28.0	23.4	29.2	25.8	28.7
入学者選抜	38.5	39.9	39.1	37.5	43.8	35.3	25.9	41.9**	34.2	43.1*
教育課程	47.5	47.3	52.3	47.2	50.3	46.2	32.5	51.7***	47.3	49.4
教育内容	56.7	61.7	50.0	54.2	60.6	54.3	41.0	61.2***	50.4	62.6**
教育方法	46.8	52.2	47.8	42.7	51.3	44.1	35.0	50.4**	42.1	52.8**
学習支援	43.1	43.7	44.1	42.3	45.4	42.2	28.2	47.3***	40.0	46.2

(出所)筆者作成

(注)*** p<0.001, ** p<0.01, * p<0.05

基礎としながら、他のアプローチが導入されることにより評価方法の多様化が進んでいるという実相が示された。8つあるいずれの評価領域でも、自律型アプローチの採用率が最も高く、教育の是非を判断する際には所属教員の専門性や見識を尊重すべきであるという発想が、確実に存在していると言える。

評価方法の多様化は、そうした発想の再検討を迫るものとして捉えられるが、必ずしもすべての評価領域において、それが均一に起こっているわけではない。標準化型アプローチの採用に表れる同等性・通用性を重視する傾向が領域を問わず見られる一方で、幾つかの領域で利用率が3割にも満たない評価方法も存在する。

応需型の評価方法は、教員構成や入学者選抜方法の評価では採用率が低く、そこでは学生の要求に応えるべきだという考え方が浸透していないと言える。入学者選抜方法の評価では、現在の学生ではなく未来の学生つまり高校生の意向を取り入れているという見方も可能だが、教員構成に関しては、学生にその良否を尋ねることは稀であると結論できるだろう。教員の職務は教育だけではないことを考慮すると当然の結果ともいえるが、教育内容や教育方法の評価では多くの学部が学生の意見に配慮しており、この隔たりを埋めるための取り組みが、近く義務化されるFD活動なのかもしれない。

もう一点、採用率の低さが顕著なのは、施設・設備や教員構成の評価における成果準拠型アプローチである。これらの領域は教育成果との繋がりを確定しづらいため、それによって是非を判断するという着想に至らないものと考えられる。

(2) 学部の属性による格差について

評価領域による格差に加え、学部の属性によっても評価方法の採用率に違いが見られた。これに関しても教育観の視点から説明できる点が幾つかある。

学部の属性で採用率に有意差がよく表れたのは、標準化型と成果準拠型の評価方法であり、前者に関しては、特に設置形態による相違が目立った。すべての領域で、私立大学での採用率が低い

のに対し、個性化の必要性が訴えられる今日においても、公設大学ではある程度の同等性が志向されていると言える。

また教育課程や教育内容の評価に関して公立大学で採用率が高くなっているのは、卒業生の職業を想定している学部で同様の傾向が見られることから、農学系・保健系の学部の多さが原因になっていると推測できる。これらの分野 農学部には獣医学科がある例も多い には、国家試験という出口の標準化を進める仕組みが備わっているため、教育そのものの評価においても標準化型が選好されると考えれば、専門分野では理系の、入試難度では上位の学部において採用率が高くなっていることも納得できる。

成果準拠型アプローチは、専門分野や設置形態でははっきりとした差が見られず、卒業生の職業が具体的に想定されている学部における採用率の高さが顕著であった。これは、就職状況や資格の取得により適切な教育成果が定義しやすいことが関係していると考えられる。領域間格差の原因と総合すると、成果を示しやすく、成果と教育との因果関係が確定しやすい場合に、より多く採用されていることになる。

なお、応需型の評価方法に関しては、学部の属性による格差はあまり見られなかったが、特に学生の意見を把握する必要があるのはどのような場合か、ということを考えて、入試難度による差異が期待された。なぜならば、葛城(2006)も示しているように、入試難度が下位の大学に在籍する学生ほど、従来想定されていた学生像とは異なることが認められているからである。学生の変質が遍く起きている証左ともとれるし、あるいは、「意見を聞くに値する学生」という教員の意識が入試難度上位の大学で働き、仮説に従う傾向を相殺している可能性もある。

(3) 今後の研究課題

今後着手が求められる研究課題としては、評価方法の採用状況に関する時系列変化の解明が挙げられる。

今回の分析からは、評価方法の多様化の実態が明らかとなったが、これは大学改革や第三者評価

の影響下において急変している過渡期のそれを描出しているとも言える。従って、評価方法の選択は、必ずしも背後にある教育観を了承した上で、意識的に行われているとは限らない。

この事実は裏を返せば、評価領域や学部の属性による採用率の格差を、意図的に是正する余地が存在するという点でもである。例えば、「教育の妥当性はその成果によって判断されるべきである」という教育観は、現状では成果の定義がしやすい学部において配慮される傾向にあるが、5年10年経った後には、学部の属性を問わず重要性が認識され、成果準拠型アプローチが普及しているかもしれない。

時系列的变化を追うことにより、単にその時々の実態が判明するだけでなく、教育観の変遷を知るための一指標を見出すこともできると考えられる。

注記

- 1 各機関が用いる評価システムの枠組みや評価結果については、ホームページから知ることができる。アドレスはそれぞれ、大学基準協会 (<http://www.juaa.or.jp/>)、大学評価・学位授与機構 (<http://www.niad.ac.jp/>)、高等教育評価機構 (<http://www.jihe.or.jp/>)。
- 2 ここでは紙幅の関係から割愛するが、詳しくは清水 (2004) 等で論じられている。
- 3 評価の定義は論者により多様だが、本稿では、「指定および把握した情報を何らかの基準に照らすことにより、ある対象の価値を判断すること」という比較的汎用性の高い定義を採用している。また本文で述べている「根拠情報」とは、この定義に含まれる「情報」を指し、評価対象に関する情報だけでなく、評価の際に考慮される情報全般を含む。従って上記の定義は、他者の評価を参考に自己評価する、というような「評価の入子構造」を許容するものである。
- 4 自己評価の方法に関心を払った日本の先行研究としては、広島大学・大学教育研究センター (現高等教育研究開発センター) が1999年に実施した全国調査をもとに評価項目分析を行った米澤他 (2000) が挙げられる。ただしここでは、評価項目に注意が向けられており、根拠情報の質的な違いは配慮されていない。またアメリカの類似研究である Peterson and Augustine (1999) では、Student Assessment (SA) と学事に関する意思決定の関係について論じているが、根拠情報のひとつであるSAの結果を相対的に位置づけることには成功していない。
- 5 技術者教育のプログラム認定機関であるJABEEが採用する評価の枠組みは、成果準拠型アプローチのこうした特徴を念頭に置いた実例のひとつである。詳しくは、<http://www.jabee.org/OpenHomePage/q&a0204-0509.htm>参照。
- 6 初等教育での実践例に関しては串本・渡辺 (2006) を、大学教育の文脈での議論としては串本 (2006) を参照されたい。
- 7 ISSAは自己評価全般に関する調査ではなく、学生についての情報把握 (student assessment) を対象としており、本論の関心と厳密に一致するわけではない。しかし全米規模の当該テーマに関する画期的調査であり、質問項目も参考になるものが少なくない。同調査については Peterson et al. (1999) を参照されたい。また前掲 Peterson and Augustine (2000) は、同調査の結果を分析したものである。
- 8 本調査は、21世紀COEプログラム「21世紀型高等教育システム構築と質的保証」プロジェクトの一環である「大学の組織改革についての調査」と同時に実施された。
- 9 質問紙においては、標準化型と成果準拠型に対応する根拠情報を、複数の項目に分けて尋ねた。前者には「学外の専門家による見解」、「評価機関による評価結果」、および「他大学・他学部の現状」が、後者には「卒業生の進路」、GPAなどの「学内指標における成績」、国家試験の合否などの「学外指標における成績」、および「教育成果に関する意見」が含まれる。上記ふたつの評価方法に関しては、対応する根拠情報をいずれかひとつでも利用している場合に、当該の評価方法を採用しているものとした。
- 10 文系学部には人文系・社会系・教員養成系およびそれらの組み合わせによる学際系、理系学部には理学系・農学系・工学系・保健系およびそれらの組み合わせによる学際系が含まれる。職業の想定に関しては、本調査における「学生の進路として、具体的な職業を想定している」という問いにおいて、「当てはまる」か「ある程度当てはまる」と答えた学部は「肯定」、「あまり当てはまらない」か「当てはまらない」と答えた学部は「否定」とした。入試難度に

は「第4回ベネッセ駿台共催記述模試」(2005年10月開催)の結果による合格率60%以上の偏差値を用い(朝日新聞社出版本部「大学」編集室2006、参照)、サンプルがほぼ二分される53以下を「下位」、54以上を「上位」とした。

参考文献

- 朝日新聞社出版本部「大学」編集室(2006)『2007年版 大学ランキング』、朝日新聞社
- 大山泰宏(2001)「大学教育評価の課題と展望」『京都大学 高等教育研究』、7: 37-55
- 串本剛(2006)「大学教育におけるプログラム評価の現状と課題 - 教育成果を根拠とした形成的評価の確立を目指して - 」、『大学論集』、37: 265-276
- 串本剛・渡辺明美(2006)「評価インターン活動報告書 - 広島県尾道市立栗原北小学校における学校評価表の開発 - 」(日本評価学会「学生インターン出前サービス」事業報告書)
- 葛城浩一(2006)「第4章 大学教育の成果としての性格・態度特性の変容」、広島大学・高等教育研究開発センター編『学生からみた大学教育の質 - 授業評価からプログラム評価へ - 』(COEシリーズ18)、55-70
- 清水一彦(2004)「大学評価の体系」、山野井敦徳・清水一彦編著『大学評価の展開』、東信堂、27-52
- 関正夫(1989)「教育評価の原理と方法に関する一考察」、『大学論集』、20: 3-31
- セイモア(館昭・森利枝訳2000)『大学個性化の戦略 高等教育のTQM』、玉川大学出版部
- 大学基準協会(2000)「大学評価の新たな地平を切り拓く(提言)」
- ドラッガー・スターン編著(田中弥生監訳2000)『非営利組織の成果重視マネジメント』、ダイヤモンド社
- 日本私立大学連盟 学生部会(2000)『ユニバーサル化時代の私立大学 そのクライアントの期待と要望』、開成出版
- 広島大学・大学教育研究センター編(1999)『大学改革と市場原理 第27回(1998年度)研究員集会の記録』(高等教育研究叢書 56)
- 牟田博光(2001)「革新と透明性向上のために - シンポジウム: 各界における評価の今日的課題と展望」、『日本評価研究』、1(1): 11-24
- 文部科学省高等教育局大学振興課(2006)「大学における教育内容等の改革状況について」(http://www.mext.go.jp/b_menu/houdou/18/06/06060504/001.pdf、2006/07/01アクセス)
- 米澤彰純・村澤昌崇・作田良三(2000)「日本の大学評価システムの構造と機能 - 自己点検・評価が生み出したもの」、『高等教育研究』、3: 173-192
- リースマン(喜多村和之ほか訳1986)『リースマン 高等教育論』、玉川大学出版部
- Barr, R. and Tagg J. (1995). From Teaching to Learning: a new paradigm for undergraduate education. *Change*, November/December, 13-25.
- Conrad, C. and Wilson, R. (1985). *Academic Program Reviews: institutional approaches, expectations, and controversies*. ASHE-ERIC Higher Education Reports, 5. Washington, DC: Association for the Study of Higher Education.
- Davies, S. (2002). Marketing in Higher Education: matching promises and reality to expectations. In OECD *Responding to Student Expectations*. 103-114.
- Eisner, E. (2003). Educational Connoisseurship and Educational Criticism: an arts-based approach to educational evaluation. In T. Kellaghan, and D. Stufflebeam (Eds.), *International Handbook of Educational Evaluation*. Dordrecht: Kluwer Academic Publisher, 153-166.
- Fitzpatrick, J., Sanders, J., and Worthen, B. (2004). *Program Evaluation: alternative approaches and practical guidelines*. Boston: Allyn and Bacon.
- Gardner, D. (1977). Five Evaluation Frameworks: implications for decision making in higher education. *Journal of Higher Education*, 48 (5), 571-593.
- Garvin, D. (1988). *Managing Quality*. New York: The Free Press.
- House, E. (1983). Assumptions Underlining Evaluation Models. In G. F. Madaus, M. Scriven, and D. L. Stufflebeam (Eds.), *Evaluation Models: viewpoints on educational and human services evaluation*. Boston: Kluwer-Nijhoff, 45-64.
- Huba, M. and Freed, J. (2000). *Learner-Centered Assessment on College Campuses: shifting the focus from teaching to learning*. Boston: Allyn and Bacon.
- Palomba, C. and Banta, T. (Eds.) (1999). *Assessment*

- Essentials*. San Francisco, Calif.: Jossey-Bass.
- Peterson, M. and Augustine, C. (2000). Organizational Practices Enhancing the Influence of Student Assessment Information in Academic Decisions. *Research in Higher Education*, 41 (1), 21-52.
- Peterson, M., Einarson, M., Augustine, C., and Vaughan, D. (1999). *Institutional Support for Student Assessment: methodology and results of a national survey*. Stanford, Calif.: National Center for Postsecondary Improvement.
- Rogers, P., Hacsí, T., Petrosino, A., and Huebner, T. (Eds.) (2000). *Program Theory in Evaluation: challenges and opportunities*. (New Directions for Evaluation, 87) San Francisco, Calif.: Jossey-Bass.
- Stark, J. and Lattuca L. (1997). *Shaping the College Curriculum: academic plan in action*. Boston: Allyn and Bacon.
- Stufflebeam, D. (2001). *Evaluation Models*. (New Directions for Evaluation, 89) San Francisco, Calif.: Jossey-Bass.
- Trigwell, K. and Prosser, M. (1996). Changing Approaches to Teaching: a relational perspective. *Studies in Higher Education*, 21 (3), 275-284.
- Tyler, R. (1949). *Basic Principles of Curriculum and Instruction*. Chicago: The University of Chicago Press.

(2007.2.1受理)

Diversification of Approaches to the Self-Evaluation of Undergraduate Education

Takeshi Kushimoto

Hiroshima University Graduate School of Education
Japan Society for the Promotion of Science Research Fellow
t.kushimoto@hiroshima-u.ac.jp

Abstract

The purpose of this study is proposing four evaluation methods and demonstrating to what degree they are adopted in self-evaluations of undergraduate education in Japan. In the first part, four evaluation approaches are introduced in relation to the ideas of undergraduate education. They are named as Autonomy-Based Approach, Demand-Based Approach, Standard-Based Approach, and Outcomes-Based Approach. In the latter part of this article, the data from a national survey of 1,871 faculty members in Japanese universities show that to what extent those four methods are used in practice. As a result of the analysis, it is found that the degree of adoption depends on the evaluation focuses and traits of the faculty. This fact is discussed in consideration of diversifying idea of undergraduate education.

Keywords

undergraduate education, self-evaluation, evaluation methods,
diversification, national survey

日本評価学会第7回全国大会
開催報告

1. 第7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 「都道府県の行政評価 この10年の実践総括」 | 梅田 次郎（JMAC構造改革推進セクター） |
| 共通論題 「教育組織の評価」 | 長尾 眞文（広島大学） |
| 共通論題 「途上国の評価キャパシティ・デベロップメント
支援に向けた取り組みの現状と課題」 | 三輪 徳子（国際協力機構）
源 由理子（国際開発コンサルタント） |
| 共通論題 「市民参加型評価」 | 岩淵 公二（政策21） |

3. 自由論題セッション報告

- | | |
|-------------------------------|---------------------|
| 自由論題 「評価手法1」 | 三好 皓一（立命館アジア太平洋大学） |
| 自由論題 「評価手法2」 | 橋本ヒロ子（十文字学園女子大学） |
| 自由論題 「保健1」 | 大島 巖（日本社会事業大学） |
| 自由論題 「評価手法3」 | 齋藤 貴浩（大学評価・学位授与機構） |
| 自由論題 「評価手法4」 | 渡辺 博（東洋エンジニアリング（株）） |
| 自由論題 「保健2」 | 梅内 拓生（国際医療福祉大学） |
| 自由論題 「行政評価/マニフェスト」 | 山谷 清志（同志社大学） |
| 自由論題 「環境」 | 長谷川 弘（広島修道大学） |
| 自由論題 「教育評価」 | 橋本 昭彦（国立教育政策研究所） |
| 自由論題 「公共財政」 | 林 薫（文教大学） |
| 特別セッション「評価インターンによる調査報告（中間報告）」 | 源 由理子（国際開発コンサルタント） |

日本評価学会第7回全国大会
開催のご報告とお礼

2006年12月2-3日開催（東洋大学）の日本評価学会第7回全国大会には170名を超える多くの方にご出席いただき誠にありがとうございました。会員各位の日頃の研究や実践活動の報告を基に、評価研究者、実務者の間の経験、情報、知識を共有化する場として、皆様にとって有益な機会となりましたら幸甚に存じます。多くの方のご尽力により本大会を開催することができましたこと、心より感謝申し上げます。今後とも当学会の活動にご高配賜りますようお願い申し上げます。

第7回全国大会

実行委員会委員長

坂元 浩一

プログラム委員会委員長

牟田 博光

日本評価学会第7回全国大会
「地域・現場に根ざした評価」(於：東洋大学)

プログラム実績

第1日：2006年12月2日(土)

9:00-9:30	受付		
午前の部 10:00-11:30	自由論題 評価手法1 (三好皓一)	自由論題 評価手法2 (橋本ヒロ子)	自由論題 保健1 (大島巖)
11:30-12:45	お昼休み(理事会 11:35-12:35)		
午後の部 12:45-14:15	自由論題 評価手法3 (齋藤貴浩)	自由論題 評価手法4 (渡辺博)	自由論題 保健2 (梅内拓生)
午後の部 14:15-16:30	共通論題 都道府県の行政評価 - この10年の実践総括 (梅田次郎)		共通論題 教育組織の評価 (長尾真文)
16:45-17:45	総会		
18:00-20:00	懇親会		

第2日：2006年12月3日(日)

9:30-10:00	受付		
午前の部 10:00-12:00	自由論題 行政評価/マニフェスト (山谷清志)	自由論題 環境 (長谷川弘)	
12:00-13:15	お昼休み		
午後の部 13:15-15:15	共通論題 途上国の評価キャパシティ・デベロップメント - 支援に向けた取り組みの現状と課題 - (三輪徳子、源由理子)		共通論題 市民参加型評価 (岩淵公二)
午後の部 15:30-17:00	自由論題 教育評価 (橋本昭彦)	自由論題 公共財政 (林薫)	特別セッション 評価インターンによる調査報告 (中間報告) (源由理子)

()は座長です。

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。

12月2日(土) 午前の部 10:00-11:30

自由論題		評価手法1	座長	三好 皓一	立命館アジア太平洋大学
10:00-10:28	評価における価値観の問題: スクリヴェンの評価論理の探求		佐々木 亮 クリス・コーリン	ウェスタン・ミシガン大学評価研究所 ウェスタン・ミシガン大学評価研究所	
10:30-10:58	貿易能力の形成援助策に関する評価のメタ分析		高橋 与志 松岡 俊二 大野 敦	広島大学 広島大学 広島大学	
11:00-11:28	二次評価における評価視点の分析		牟田 博光 菊田 怜子	東京工業大学 東京工業大学	
自由論題		評価手法2	座長	橋本ヒロ子	十文字学園女子大学
10:00-10:28	エンパワメント評価モデル: 成果三類型モデルに関する考察		藤掛 洋子	東京家政学院大学	
10:30-10:58	ジェンダー評価について		青木 憲代	アイ・シー・ネット(株)	
11:00-11:28	ソーシャルキャピタルに基づく住民組織の「手探り」評価方法の考え方		渡辺 淳一	(財)国際開発センター	
自由論題		保健1	座長	大島 巖	日本社会事業大学
10:00-10:28	地域健康資産形成に関するアプローチの評価 - 老人介護病院経営の視点から -		佐藤 辰夫 梅内 拓生	医療法人社団葵会青梅今井病院 国際医療福祉大学大学院	
10:30-10:58	地域保健資産形成へのアプローチの検討 - 医療法人と行政の協力の視点から -		真壁 道夫 羽根田 潔 梅内 拓生	医療法人医徳会 医療法人医徳会 国際医療福祉大学	
11:00-11:28	適切な医薬品使用に関する考察 - 地域ケア薬局に関するアンケート調査 - (薬局薬剤師と薬学部学生との比較)		梅原 貞臣 野口 隆志 高他 武始 梅内 拓生	国際医療福祉大学 国際医療福祉大学 国際医療福祉大学 国際医療福祉大学	

12月2日(土) 午後の部 12:45-14:15

自由論題		評価手法3	座長	齋藤 貴浩	大学評価・学位授与機構
12:45-13:13	要求水準の変化を考慮した評価測定		日野 哲也 牟田 博光	東京工業大学大学院 東京工業大学大学院	
13:15-13:43	事業分野別社会資本の整備水準と労働移動 - 2財一般均衡モデルをもとにした最適性の評価 -		鈴木 美和	東北大学	
13:45-14:13	開発と利益の2つの視点からのODA評価についての考察 - 標準化した効果単位を使い、インパクトを定量化した場合 -		佐藤由利子	東京工業大学	
自由論題		評価手法4	座長	渡辺 博	東洋エンジニアリング(株)
12:45-13:13	死亡交通事故対策の評価について - 北海道の急激な交通事故死者数減少の要因について -		浅野 基樹	土木研究所・寒地土木研究所	
13:15-13:43	道路満足度を向上するための信頼度向上施策について		大島 章嘉	市民満足学会(株)ワード研究所	
13:45-14:13	AOTS研修の事業評価システムと品質保証体系		神田 久史	海外技術者研修協会	
自由論題		保健2	座長	梅内 拓生	国際医療福祉大学
12:45-13:13	サービス特性からみたPFI事業の評価手法 - 高知医療センターのケーススタディ -		山下 興一 渡里 司 松岡 俊二	広島大学大学院 中国地方総合研究センター 広島大学大学院	
13:15-13:43	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() - 国別の乳児死亡率と一人当たりのGDPのベキ法則関係上の世界レベルでの位置変化と国自身の経時的変化の評価学的意味の検討 -		正木 朋也 梅内 拓生	国際医療福祉大学大学院 国際医療福祉大学大学院	
13:45-14:13	精神保健福祉領域におけるプログラム評価の近年の発展と課題		大島 巖	日本社会事業大学	

12月2日(土) 午後の部 14:30-16:30			
共通論題	都道府県の行政評価 - この10年の実践総括	座長 梅田 次郎 討論者 山谷 清志	JMAC構造改革推進セクター 同志社大学
1	都道府県の行政評価、この10年の歩みと今後の課題		谷口 敏彦 行政経営アナリスト
2	都道府県の行政評価、この10年間の問題点 - 定量評価の陥穽 -		小野 達也 鳥取大学
3	事務事業評価の10年: 正確な評価を目指して		窪田 好男 神戸学院大学
共通論題	教育組織の評価	座長 長尾 眞文	広島大学
1	学校評価の現状と課題について		岸本 哲哉 文部科学省
2	経営改善に資する学校評価の実際		石田 謙豪 尾道市立栗原北小学校
3	大学組織の評価: 質的向上と説明責任の支援・促進		齋藤 貴浩 大学評価・学位授与機構
4	認証評価制度における自己点検評価者としての評価 アプローチ - インターベンション評価 -		西出 順郎 琉球大学

12月2日(土) 総会 16:45-17:45

12月2日(土) 懇親会 18:00-20:00

12月3日(日) 午前の部 10:00-12:00			
自由論題	行政評価/マニフェスト	座長 山谷 清志	同志社大学
10:00-10:28	評価制度と自治体職員の意識変化		横山麻季子 筑波大学大学院 鶴田 芳貴 筑波大学大学院
10:30-10:58	韓国における市民社会の公約評価活動		河東 賢 慶応義塾大学
11:00-11:28	行政評価と事業仕分		佐藤 章 大津役所
11:30-11:58	地方自治体の行政評価: 10年間の実践を診断する		田中 啓 静岡文化芸術大学
自由論題	環境	座長 長谷川 弘	広島修道大学
11:30-11:58	過去20年にわたる日本の対中環境協力について		廣野 良吉 成蹊大学
10:30-10:58	社会的環境管理能力の形成プロセスに係る因果構造 の分析: 都市大気汚染対策を事例にして		村上 一真 広島大学 松岡 俊二 広島大学
11:00-11:28	企業経営と地域経済環境資産形成へのアプローチ 評価(企業経営の視点から)		築島 武彦 (株)日本ギャラクシーエンジニアリング 梅内 拓生 国際医療福祉大学大学院
11:30-11:58	日本における環境教育の評価の現状と課題: ESDの評価を視野に入れて		藪並 郁子 立教大学大学院 阿部 治 立教大学

12月3日(日) 午後の部 13:15-15:15

共通論題	途上国の評価キャパシティ・デベロップメント - 支援に向けた取り組みの現状と課題 -	座長 三輪 徳子 座長 源 由理子	国際協力機構 国際開発コンサルタント
1	開発途上国における評価能力向上支援の現状と課題 - DACによる主要ドナーのECD支援状況調査結果より -		源 由理子 国際開発コンサルタント 三輪 徳子 国際協力機構 藤本 真美 外務省
2	評価キャパシティ開発(ECD)へのJBICの取組		和田 義郎 国際協力銀行
3	開発途上国のECDへのニーズ: 評価セミナーと合同評価の経験より		原口 孝子 (株)国際開発アソシエイツ
4	開発途上国の評価フィードバックとキャパシティ・デベロップメント		湊直 信 (財)国際開発高等教育機構
共通論題	市民参加型評価	座長 岩淵 公二 討論者 窪田 好男	政策21 神戸学院大学
1	市民参加と評価の仕組み		今川 晃 同志社大学
2	参加型評価の類型と機能: 評価専門家の役割の視点から -		藤田 伸子 (財)国際開発高等教育機構
3	協働事業を対象とした市民参加型評価を実践して		高橋 敏彦 いわてNPO-NETサポート

12月3日(日) 午後の部 15:30-17:00			
自由論題	教育評価	座長	橋本 昭彦 国立教育政策研究所
15:30-15:58	義務教育課程における学校評価導入に関する考察		小久保純一 経営コンサルタント 教育コンサルタント 前三重県朝日町教育委員会教育長
16:00-16:28	学校評価における第三者評価の意義とその課題		森田 祐司 武久 顕也 村上 恵美 植田みどり 監査法人トーマツ 監査法人トーマツ 監査法人トーマツ 国立教育政策研究所
16:30-16:58	モンゴル農村部における教育の質的向上：国連・ユネスコプロジェクト評価結果からの政策提言		山口しのぶ 矢野 智子 東京工業大学 ユネスコ北京事務所
自由論題	公共財政	座長	林 薫 文教大学
15:30-15:58	途上国の公共財政制度と財政執行の評価方法2：PEFA業績測定枠組みへの業績評価報告書の考案		上野 宏 南山大学
16:00-16:28	「公共財政管理」の枠組みによる公的債務管理返済能力評価（モルディブの事例）【発表キャンセル】		尾崎 隆夫 狩野伊知郎 黒岩 宙司 東京大学大学院 （社）海外コンサルティング企業協会 東京大学大学院
16:30-16:58	貧困国における「一般財政支援」の評価		山下 道子 （財）国際開発センター
特別セッション	評価インターンによる調査報告（中間報告）	座長	源 由理子 国際開発コンサルタント
15:30-15:58	市民フォーラム21・NPOセンターの組織評価		松尾 武司 立教大学大学院
16:00-16:28	ロジックモデルによる学校評価の改善		竹下 雅美 学習院女子大学大学院
16:30-16:58	授業改善を目的とした生徒による授業評価システムの開発と実践		山崎 その 市野 緑 丸山 真司 同志社大学大学院 東京工業大学大学院 東京工業大学大学院

共通論題セッション報告

共通論題セッション 「都道府県の行政評価 この10年の実践総括」

座長 梅田 次郎（JMAC構造改革推進セクター）

討論者 山谷 清志（同志社大学）

1996年の三重県における本格導入から現在までにほとんどの都道府県において行政評価が導入されており、良くも悪くも評価の実績量としては都道府県が最も多い。本セッションの目的は、この10年間の実践について動向をまとめ、またその問題点を摘出しながら、今後の展望を見極めようとするものである。

第1報告者の谷口敏彦会員（行政経営アナリスト）は「この10年の歩みと今後の課題」と題して、各種のアンケート結果を基に行政評価の実施状況と成果の活用方法、評価の実施による効果などを要約するとともに、先行的な取組を行っている4県の取組の概要を紹介し、今後の課題は評価結果をどのように活用するかであり、行政評価は実施する段階から行政経営にどう生かすかの段階に移ってきたと報告された。

第2報告者の小野達也会員（鳥取大学）は「この10年間の問題点 定量評価の陥穽」と題して、評価指標やアンケート集計値の利用など定量的な方法が重視される傾向は評価の客観性を高めるという観点から望ましいと考えられるが、具体的にみていくと数々の課題があるといわざるを得ないとし、行政評価で用いられる定量的な方法に見られる問題の所在を明らかにし、それが評価の過程や結果の活用に及ぼす影響について報告された。

第3報告者の窪田好男会員（神戸学院大学）は「事務事業評価の10年：正確な評価を目指して」と題して、評価の到達点と問題点を確認した後、解決策として議論されている評価の客観性の向上には難点が多いことを指摘し、政策体系のごく一部だけでも正確な評価を十分なコストをかけた外部評価によって行い、自己評価と対照し、不正確な自己評価を行っていた評価者が「罰」を受けるようにすることで評価全体の正確さを向上させることを提言された。

これらを受けて討論者の山谷清志会員（同志社大学）からは、行政評価が行政のマネジメントに役立つことと住民にわかりやすいこととの連立の困難性などについてふれながら「分かりやすさ」とは何だろうとの問いかけがなされ、各報告者からそれぞれ見解が述べられた。

会場からも活発な質疑や意見が発表され、実務と研究の協働の場である日本評価学会ならではの議論が展開された。

最後に私から、当学会の会員名簿をみると省庁の職員や自治体の職員が極めて少ないことにかんがみ、今回は都道府県を扱ったが、国、市町村の評価実績も積み重なってきているので、それぞれをテーマにして実務の参考にもなるようなセッションを今後鋭意計画すべきであるとの提言をさせていただき終了しました。

共通論題セッション 「教育組織の評価」

座長 長尾 眞文（広島大学）

一般に組織にとって、評価は 説明責任の遂行、 事業・運営の改善、 効率的・効果的な意思決定のための知識管理の3課題への対応に役立つとされる。教育組織の場合に、どのような目的が設定され、実際

に誰がどのような方法を用いて評価を行うかについては未だ定見はなく、試行錯誤を重ねている段階である。そこでは、まず他の公共部門における評価と同様に、評価プロセスの透明性をいかに確保するか、評価の実施に必要な専門性をいかに調達するか、そして評価結果に対する信頼性をいかに獲得するかといった組織評価が本来的に抱える課題が浮上する。また、教育基本法の改正、教育の地方分権化や教育機関と地域や産業との関係のあり方の変化といった教育活動にとって外部的な条件の変化が教育組織の評価に多様な影響を及ぼしている。これら内部的、外部的要因の双方に配慮しつつ教育分野で変革を支援するツールとして評価が定着するには、評価の有用性について多角的に検証する必要がある。

本セッションでは、義務教育段階の教育と大学教育の的を絞り、それぞれについて内部的、外部的関係者から報告をもらった上で、教育組織にとって評価が役に立つ条件を議論した。第1報告者の岸本哲哉氏（文部科学省初等中等教育局学校評価室）は、現在初等中等教育段階での学校評価が全国的に普及しつつある現状について、その概要、法的根拠、制度上の課題、文科省の学校評価ガイドライン（2006年3月公表）の内容を中心に説明するとともに、国の視点から学校評価システムの運用とその有用性の鍵を握るとされる第三者評価び扱いについて展望を語られた。これに対して、第2報告者の石田謙豪会員（広島県尾道市立栗原北小学校）は、自らが校長を務める小学校で、どのように経営改善に資する学校評価の実現と取り組んでいるかについて、保護者の願いを取り入れた教育活動を行うことの重要性和教員参画による評価活動の推進の必要性を強調しつつ報告された。この前半の義務教育段階の学校の評価に関する質疑応答では、教育の質の保証をどのように評価するかという方法的関心が主として取り上げられ、学力テストで小中学校段階での理解度を十分反映できるか等が議論された。

セッション後半では、まず第3報告者の齋藤貴浩会員（大学評価・学位授与機構）は、大学認証組織の立場から、大学組織の評価について評価機関の側で対応すべきことと大学側で対応すべきことの双方の視点から報告され、続けて第4報告者の西出順郎会員（琉球大学）は、認証評価制度に直面した大学が自己点検評価とどう取り組むかについて、インターベンション評価の概念を使って議論された。質疑応答では、高等教育機関の評価で競争条件や組織の個性をどう評価するのか、組織評価が日本では機能するのか等が取り上げられた。

セッション全体を通して、教育組織の評価が日本における評価慣行の普及の最前線となりつつあるとの認識が共有され、その将来的な展開にとって高等専門性を備えた評価人材の育成が急務であること、そしてそれは教育組織、特に大学が自ら取り上げるべき課題であることが確認された。また、その関連で、学会が認証制度の運用等を通して果たすべき役割にも議論が及んだ。教育組織の評価は単に評価の一分野であるだけでなく、評価人材の育成にも係わることから、学会としても研究の活性化や会員の発掘に努める必要が感じられた。

共通論題セッション 「途上国の評価キャパシティ・デベロップメント 支援に向けた取り組みの現状と課題」

座長 三輪 徳子（国際協力機構）

座長 源 由理子（国際開発コンサルタント）

援助効果向上に向けた国際的な取り組みが加速する中、途上国のオーナーシップの確保や開発成果のマネジメントの観点から、途上国の評価キャパシティ・デベロップメント（ECD）がこれまで以上に注目、重視されつつある。本セッションではこうした状況をふまえ、今後の途上国ECDの効果的な支援のあり方を探ることを目的に、ECDとその支援の現状と課題について様々な角度から報告をいただき、議論を行った。

第1報告者の源由理子氏（国際開発コンサルタント）からは、DAC開発評価ネットワークによるECD支

援状況調査結果に基づいて主要ドナーのECD支援の概況が報告され、途上国側の評価に対するコミットメントの低さや制度面の脆弱さ、またドナー側における協調の欠如やドナー主導の弊害などがECDの課題として挙がっていることが示された。ついで第2報告者の和田義郎氏（JBIC）からは、MDGs達成に向けた取り組みの文脈の中でECDを捉える必要があるとしたうえで、計画官庁に対するECD支援に加えて、実際と同取り組みにあたる実施機関や地方政府等のECD支援も重要との指摘がなされた。また、第3報告者の原口孝子氏（（株）国際開発アソシエイツ）からは、途上国向け評価研修や途上国との合同評価に携わった経験に基づき、ECDのためには評価技術の習得のみでなく体制構築が必須であり、途上国における評価ニーズを高める工夫を含む取り組みが求められるとの問題提起が行われた。最後に第4報告者の湊直信氏（（財）国際開発高等教育機構）からは、評価文化の確立と幅広い関係者のECDに向けて評価学会など知のネットワークが果たす役割が論じられ、途上国における国レベル・地域レベルの評価学会設立の動きや活動が報告された。

これらの報告に対して会場からは多数の質問・意見が出され、ECDにおける途上国の主体性、評価結果フィードバックの仕組みづくり、ドナー側評価手続きの調和化の問題などについて活発な議論が行われた。技術面のみでなく制度面を含むECDへの取り組みの重要性が共有され、評価と政策とのリンク、評価文化の構築など幅広い視点からECDを捉え、途上国主導によるECDのあり方について議論が深められたことは、本セッションの成果であったと言える。

共通論題 「市民参加型評価」

座長 岩淵 公二（特定非営利活動法人政策21）
 討論者 窪田 好男（神戸学院大学）

本セッションの目的は、大会テーマである「地域・現場に根ざした評価」の重要な論点として、政策評価・行政評価における市民の役割を探ることにあつた。

近年、政策過程への市民参加が重要性を増し、市民による評価を試行する自治体が増えている。いわゆる「参加型評価」と称される評価の形態であるが、その中には評価の目的を考えるうえで重要な二つの思考系譜が含まれている。すなわち、政策決定過程への参加プロセスを重視する「市民参加」であり、もう一つは市民意見を取り入れた評価を行う手法として活用されてきた「参加型評価」である。これら二つのキーワードは、手法面では類似性を持ちつつも、その目的と関心は同一ではない。そこで、理論と実践の双方から市民参加型評価の課題と展望を議論するため、市民参加、参加型評価、事例報告の三つの観点から報告いただいた。

第一報告者の今川晃同志社大学政策学部教授は、「市民参加と評価」と題して発表し、市民参加には水平的自律的調整機能が期待されていることを指摘したうえで、多様な主体による評価の必要性を報告した。第二報告者の藤田伸子会員（国際開発高等教育機構）は、「参加型評価の類型と機能」と題して発表し、参加型評価のアプローチを目的、評価専門家の関与、ステークホルダーの関わりの方から整理したうえで、参加型評価における課題について報告した。第三報告者の高橋敏彦いわてNPO-NETサポート代表理事は、「市民参加による評価 協働事業の外部評価」と題して発表し、岩手県において今年度実践した県民参加型外部評価の事例を報告したうえで、市民参加による評価の効果と課題について報告した。報告後、討論者の窪田好男会員（神戸学院大学）が、参加する市民も含めた評価者育成の重要を提起し、フロアを交え報告者との議論を行った。

本セッションでは、テーマに多角的な視点からアプローチすべく、学会外から市民参加に詳しい研究者と市民参加による評価の実践者を招き議論したものであるが、市民参加型評価の課題と今後の展開を考えるうえで、有意義な議論ができたと考えている。

自由論題セッション報告

自由論題セッション 「評価手法1」

座長 三好 皓一（立命館アジア太平洋大学）

本セッションでは、3つの報告がなされた。

まず、佐々木会員（ミシガン大学評価研究所）から、「評価における価値観の問題：スクリヴェンの評価論理の探求」について報告が行なわれた。本報告は、スクリヴェンが長年にわたり唱導してきた「評価論理（Logic of Evaluation）」について、評価の定義、評価における価値判断、価値判断のための価値観を客観的に明らかにする作業、評価の将来像としての分野横断的モデルについて解説し、その意義を紹介するとともに、その日本の評価及び評価研究への示唆を提示するものであった。フロアーからは、価値の多様化、また、国際協力銀行の評価におけるレーティングなどに関連して質問、コメントがなされた。なお、本報告はテレビ会議システムを活用してアメリカと接続して行っており、佐々木会員の報告に先立ち本セッション参加者に対しスクリヴェン教授から挨拶があった。

次に、牟田博光会員（東京工業大学）から、「二次評価における評価視点の分析」について報告が行なわれた。本報告は、国際協力機構（JICA）の二次評価についてその概要の説明・紹介とともに、JICAが実施したプロジェクトの終了時評価報告書を対象とした国際協力専攻の大学院生による2回の二次評価の実施結果を提示するものであった。結果として、評価を繰り返すことによる評価が辛くなる傾向について、また、評価基準の定まりにくい項目、評価が容易でない項目などについて報告がなされた。フロアーからは、評価結果のばらつきの状況、1回目と2回目の評価の間隔に伴う影響の可能性、一次評価と二次評価の違いなどに関して質問、コメントがなされた。

最後に、高橋与志会員（広島大学）から、「貿易能力の形成援助策に関するメタ分析」について報告が行なわれた。本報告は、これまでに発表された援助実施機関や機関横断的な援助を対象とした主な評価報告書の比較分析を通じて評価方法としての整理を試みたものである。結果として、対象とした評価では妥当性の主要な構成要素である「必要性」「優先度」「手段としての適切性」のすべてを捕らえたものが無い点とともに、より包括的な評価基準の設定の必要性、ベンチマーク・指標の開発についての今後の課題としての重要性、また、新戦略との関連性についての研究の必要性などについての指摘がなされた。フロアーからは、比較優位との関連、また、ケースの選択方法などについて質問、コメントがなされた。

自由論題セッション 「評価手法2」

座長 橋本 ヒロ子（十文字学園女子大学）

「評価手法2」では以下のようにジェンダー評価に関する報告2点と、ソーシャルキャピタルに関する評価の報告1点、合計3点の報告があった。各報告に引き続き、質疑及び示唆に富んだコメントがされた。

1. エンパワーメント評価モデル：成果三類型モデルに関する考察

藤掛 洋子（東京家政学院大学）

藤掛氏は、南米パラグアイ共和国農村の女性たちが生活改善プロジェクトに主体的にかかわるプロセス

を1993年から2006年まで調査して、エンパワーメント評価のモデルである成果三類型モデル（以下、三類型モデル）を検討してきた。本報告では、まず、質的評価としてのエンパワーメント評価の系譜について触れ、次に三類型モデルの特徴を紹介した。その上で、MDGsの一つにも掲げられている女性のエンパワーメントという現象の可視化の可能性と課題について、中米パラグアイのプロジェクトで検証した三類型モデルのホンジェラスでの応用事例を踏まえ考察し報告した。

2. ジェンダー評価について

青木 憲代（アイ・シー・ネット（株））

本報告は、JICA、外務省等が実施したジェンダー評価の評価をして、今後のジェンダー評価に活用することを目的としている。日本ではジェンダー評価自体の数が限られていること、評価手法が確立していないことなどにより、メタ評価が困難であるという問題点を指摘し、ジェンダー評価の蓄積を提案した。

3. ソーシャルキャピタルに基づく住民組織の「手探り」評価方法の考え

渡辺 淳一（（財）国際開発センター）

渡辺氏は、ソーシャルキャピタルにおける住民組織等の現状把握をする調査手法である世界銀行のSOCAT（The Social Capital Assessment Tool）を活用して、制度的ソーシャルキャピタルの質と量の変化に焦点を当てた住民組織の評価の再検討を提案した。その際、藤掛氏の三類型累計型モデルの活用の可能性に言及した。

自由論題セッション 「保健1」

座長 大島 巖（日本社会事業大学）

本セッションでは、保健の実践現場における、新しい概念を用いた、新しい取り組みの評価に関わる3題の報告が行われた。

まず第一報告佐藤辰夫会員の「地域健康資産形成に関するアプローチの評価」では、老人医療施設（病院）に入院している高齢者の心のケアに対する病院全体での取り組みを、病院資産、特に地域健康資産という観点から検討した結果、病院職員全体のモラル向上や地域住民の心のケア向上に資することが明らかになったことが報告された。心のケアを目指した地域健康資産育成活動としては、地域小学校生徒の施設訪問、職員研修・教育活動、地域健康講話、自治体との連携など活発な活動が営まれ、その成果として医療介護サービスの改善のみならず施設維持費の節約効果も求められた。質疑では、地域健康資産という新しい概念について、その有効性が討議された。

また、第二報告眞壁道夫会員の「地域保健資産形成へのアプローチの検討 医療法人と行政の協力の視点から」では、宮城県にある医療法人が法人総収入の0.2%を投資し、地元自治体と協力して地域のプライマリヘルスケア、健康増進、住民参加、男女共同参画などの活動を行い、その成果を地域保健資産という観点から評価した。医療法人と自治体は連携して、地域保健資産の育成を進めているが、そのプロセスでは現状分析から政策の構築、情報発信、政策実践、評価と進んでいく。医療法人が、現状ではプロセスの前半部分を分担する、一見すると行政との役割が逆転する取り組みになっているが、その形態が健康政策作りにノウハウを持たない地域における保健施策構築と評価の一つの先行モデルになるものと思われる。

第三報告梅原貞臣会員の「適切な医薬品使用に関する考察 地域ケア一薬局に関するアンケート調査（薬局薬剤師と薬学部学生との比較）」では、薬局に従事する薬剤師および、これから薬剤師の資格取得を目指している薬学部学生を対象に、「政府施策の周知度」「薬局の現状」「地域ケアの薬局への問題点」

に関する自記式調査を実施し、多くの回答者が地域住民建國ケアの担い手として薬剤師の重要性を認識しながら、現在の薬局の業務の繁忙さ、政府施策の周知度の不徹底等により十分対応出来ていないことが報告された。本研究は、「適切な医薬品使用と教育（Optimal and Rationale Drug Use and Education）」という理念を検証する研究の一環として実施されているが、実証データがその検証にどのように貢献しているのかがやや不明確で、質疑の中で討議が行われた。

第一報告と第二報告は、地域健康（保健）資産という概念導入の有効性についての検討、第三報告は「適切な医薬品使用と教育」という概念の検証という新しい概念の提起と評価という点で共通する報告であった。これらの概念の有効性の評価を実証的に明らかにする課題は、とても魅力的であるが、今回の報告では必ずしもそれが十分に検証されているとは言えず、今後の研究の発展が望まれるように思われる。

自由論題セッション 「評価手法3」

座長 齋藤 貴浩（大学評価・学位授与機構）

本セッションは以下の3件の発表の後、質疑応答が行われた。

「要求水準の変化を考慮した評価測定」（日野哲也、牟田博光）では、質問紙調査等を用いてプロジェクトを評価する際に、被験者の要求水準が変化することによって真の効果が歪められている可能性と、その影響を除こうとする手法が論じられた。プロジェクトの効果を利害関係者の意識に求める場合、各個人の「判断」が効果測定の基礎となるが、その判断は必ずしも合理的ではない。評価の際にその影響を考慮しようとする試みであり、実際の分析が待たれる。

「事業分野別社会資本の整備水準と労働移動 - 2財一般均衡モデルをもとにした最適性の評価 -」（鈴木美和）では、2財一般均衡モデルを元に、日本の都道府県データから、道路・港湾などの産業基盤関連や農林水産関連の社会資本の整備水準は過大である可能性がある一方、生活環境関連や国土保全関連の社会資本の整備水準は過小であるとの結果が報告された。モデルとしての制約条件はきついが、実データをモデルに当てはめて現状の社会資本を評価した本報告は、評価への経済モデルの援用に重要な示唆を与えた。

「開発と国益の2つの視点からのODA評価についての考察 - 標準化した効果単位を使い、インパクトを定量化した場合 -」（佐藤由利子）では、日本の留学生政策に関して「人材養成」と「親日家養成」という二つの目的から、その成果の定量的評価が報告された。タイとインドネシアを対象として実データからその結果を比較するとともに、その際の分析方法として成果の標準化によってインパクトの比較を可能とする方法が提案された。

これらの発表に対し各々質疑応答が行われた。評価手法は万能ではなく、必ず何かしらの制約条件や分析上の限界が存在する。新しい評価手法や評価モデルの開発は重要であり、それが評価の領域を拡大し、またより精緻な評価を可能とする。しかし、同時に、実用に際しては、その限界や制約を十分に考慮した上での使用が評価者に求められるということが再認識された。

自由論題セッション 「評価手法4」

座長 渡辺 博（東洋エンジニアリング（株））

本セッションは以下の3件の発表の後、質疑応答があった。

「死亡交通事故対策の評価について - 北海道の急激な交通事故死者数減少の要因について -」（浅野基樹）では、北海道の交通事故死者数が平成12年の548人から平成17年の302人へこの5年間に約45%の急激な減少を見、平成17年はそれまで13年続いてきた都道府県別交通事故死者数ワースト1を14年振りに返上し、

ワースト4位となったこと、若年層の死亡交通事故の現状と若者の行動変容、自動車の安全性向上、道路交通法改正、道路整備、および交通安全運動と取り締まり等の面から交通事故死者数の減少要因について報告された。

「道路満足度を向上するための信頼度向上施策について」(大島章嘉)では、直轄国道の国民満足度を向上させようとする、不満の理由の解消が第一義的に求められるが、信頼度の向上が重要であることが明らかとなっていること、満足度と信頼度の関係、不満の理由、不信の理由の把握、施策の方向性などについて意識調査をもとにした評価が行われていること、この意識調査では調査主催者の偏見や恣意性を排除する工夫がされており利用者・回答者の本音の意見が回答に十分に反映されるようになっていることが報告された。

「AOTS研修の事業評価システムと品質保証体系」(神田久史)では、財団法人海外技術者研修協会(AOTS)で、研修の効果・効率性を高めるべく導入した事業評価システムの一般研修での運用に際して、従来の研修目標や課題を整理し、研修内容を見直す必要性が高まってを受け、3年前に一般研修の標準化計画を開始したこと、現在では研修内容の更なる改善に努めると共に、研修に対する品質保証体系の確立を目指し活動を行っていること、品質保証体系の構築の経緯と内容、並びにその構築過程において役割を変化させつつある事業評価システムについて報告された。

これらの発表について内容に関する質疑応答が行われた。特にAOTS研修事業評価システムに関しては、研修事業を製造業やサービス業に見られるような継続的な製品やサービスの生産過程に類似しているものと捉え、品質保証体系の一部と位置づけて評価システムを適用しており、極めてユニークなものであり、職業訓練評価、教育評価等の参考として示唆的なものであった。

自由論題セッション 「保健2」

座長 梅内 拓生(国際医療福祉大学大学院)

1. サービス特性からみたPFI事業の評価手法 - 高知医療センターのケーススタディ -

山下 興一(広島大学大学院)、渡里 司(中国地方総合研究センター)、松岡 俊二(広島大学大学院)
財政収支の悪化を背景に、地方自治体において官民協働(Public Private Partnership, PPP)に基づく事業が盛んに行われており、そのなかで民間資金等の活用による公共施設等の整備(Public Finance Initiative, PFI)のひとつとして高知医療センターPFI事業を選び、その評価手法に関する研究である。研究目的の設定はタイムリーなものであり、評価項目の設定も広く用いられている五項目(目標達成度、効率性、インパクト、持続性、妥当性)を用いて行っている。医療サービス機関の評価においては、財政面改善のみならず、医療サービスの質の向上が大切である。この点もValue For Money(VFM)として研究の視野に入っており、このことから、VEMと医療サービスの質の向上の評価に関する今後の研究の進展が期待される。

2. 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() - 国別の乳児死亡率と一人当たりのGDPのベキ法則関係上の世界レベルでの位置関係と国自身の経時的変化の評価の意味の検討 -

正木 朋也(国際医療福祉大学大学院)、梅内 拓生(国際医療福祉大学大学院)
乳児死亡率(IMR)と一人当たりのGDP(GDP per Capita)の関係がベキ法則に従っていることを示し、これらの関係を世界レベルと国自身のレベルで検討し、各国のベキ法則の数式上の位置の変化の動向を検討している。このことから、世界レベルの動向としては、過去半世紀の経済発展に見られるように多くの国でIMRとGDP per Capitaベキ関係が一定の方向に相伴って変化しきっていたが、ここ十年来先進国の動向

に従来のものとは異なる変化を示唆する兆候が観察されてきており、その分析の必要性を提示している。さらに国レベルの動向の検討から、各国自身の変化の軌道を整理して、軌道とIMRとの関係を政策的対応と結びつけることによって、これが政策評価に繋がる可能性を提示している。今後の研究の進展を期待したい。

3. 精神保健福祉領域におけるプログラム評価の近年の発展と課題

大島 巖（日本社会事業大学社会学福祉部）

精神保健福祉領域のプログラム評価では、科学的根拠に基づくEBM（Evidence-Based Medicine）や科学的根拠に基づく実践プログラム（EBP; Evidence-Based Practice）が注目されており、ここでは、プログラム評価方法論における実践的な発展が展開されている。大島はこの新しい動向を整理して提示するとともに、アウトカムに貢献するプログラムモデルの構築とFidelity評価法の発展、サービス普及研究の発展とEBPツールキットプロジェクトの意義という観点から、プログラム評価の方法的発展と今後の課題に検討を加えている。留意すべき点はEBMに基づく評価と異なりEBPの評価ではEBPにおけるプログラムとのサービスギャップである。EBMへのFidelityを視野に入れたEBP実践とその評価の進展は今後の研究に期待するものがある。

自由論題セッション 「行政評価/マニフェスト」

座長 山谷 清志（同志社大学）

本セッションは、横山麻季子・鶴田芳貴「評価制度と自治体職員の意識変化」（筑波大学大学院）河東賢「韓国における市民社会の公約評価活動」（慶応義塾大学）佐藤章「行政評価と事業仕分」（大津市役所）田中啓「地方自治体の行政評価：10年間の実践を診断する」（静岡文化芸術大学）の4報告で行われた。行政評価やマニフェスト、事業仕分けという一見脈絡の無い構成であるが、実は評価制度の定着とその課題という隠れた共通論題があった。

第1報告「評価制度と自治体職員の意識変化」では地方自治体による評価制度の導入やその活用が職員の意識変化に与える影響について実証分析を行なっている。この報告では、評価が政策・施策・事務事業の策定や見直し、すべての事務事業を対象として評価を長期にわたって行うこと、記入済み評価シートの公開などが、職員の意識変化につながることを明らかにした。

第2報告「韓国における市民社会の公約評価活動」では、2006年2月から韓国ではじまった地方選挙に向けてマニフェスト運動をとりあげている。韓国マニフェスト運動の特徴は、市民社会による各候補者の公約評価活動が活発に行われたことであり、報告によれば531マニフェスト政策選挙推進本部、2006地方選挙市民連帯などの各市民団体がそれぞれの評価団、評価基準、評価方式に基づき、評価活動を行っている。こうした経緯をふまえ、報告では2つの市民団体の評価経緯、評価対象・基準・方式、評価公表の方法を検討して、今回のその活動の特徴と限界を分析している。

第3報告「行政評価と事業仕分」は、地方自治体で近年取り入れ始められている事業仕分けについて、それが十分に吟味され、検討されて導入されているのかどうか疑わしい事実を指摘している。行政評価を補完するという点では意味があるが、事業を廃止する際の立論に主観が入り込む、区分を担当する人びとの立場に正統性が薄いといった致命的問題が、この方法を一時の流行に終わらせるのではないかという危惧を示している。

第4報告「地方自治体の行政評価：10年間の実践を診断する」は、過去10年間に日本の自治体で広く普及した行政評価・政策評価が、「それらを実践した結果どのような成果が上がっているのか」について明らかにしようとしている。そこで自治体に対する独自の調査を実施、行政評価を実践した成果を把握する

ことを試み、また行政評価への取り組み方法と評価制度のパフォーマンスとの関係についても分析を行っている。以上の分析を踏まえ、この報告では自治体の行政評価が抱える課題とめざすべき今後の方向性を示している。

会場からは個々の報告にさまざまな意見が出された。そのすべてをここで紹介する紙幅はないが、座長の全体的な印象としては、第1報告から第3報告までは、種々の評価を「誰が」「どうした場で」「どのように活用するのか」、そして「そのための制度設計をどのようにするのか」という質問が出されていた。また、第4報告に関しては共通論題「都道府県の行政評価」との関連もあり、地方自治体に対するアンケート調査方法の再検討、および制度のパフォーマンスを規定する要因の分析、の2点が必要ではないかという声会場から上がった。なお、座長はこのセッションの終わりに「自由論題報告者は『日本評価研究』に積極的に挑戦して、報告を学会の重要な財産にしてほしい」と発言し、この自由論題を締めくくっている。

自由論題セッション 「環境」

座長 長谷川 弘（広島修道大学）

環境分野における評価手法とその成果をテーマとした本セッションでは、次の4報告の発表とそれに続く質疑応答や意見交換が行われた。

「過去20年にわたる日本の対中環境協力について」(廣野良吉)は、数多くの関連資料や報告者の現地調査を踏まえ、環境分野におけるこれまでの対中協力の歴史的経緯、環境政策・人づくり・制度整備等への成果を分析し、中国の環境問題解決はもとより同国の国際的地位向上にも少なからず貢献したとの評価を与えている。また、日本の経験に基づく協力が求められる今後の課題として、中国の環境関連組織の強化や意識改革、更には民間活力導入に向けての環境ガバナンスの改善等が指摘された。

「社会的環境管理能力の形成プロセスに係る因果構造の分析：都市大気汚染対策を事例にして」(村上一真、松岡俊二)では、横浜市、名古屋市及び大阪市の大気環境行政活動を対象に緻密なデータ解析・統計分析を通じ、アクター間の役割の代替を経て社会的環境管理能力が形成されること、行政の環境管理能力の向上が市民・企業のそれも向上させると因果関係が存在することが実証された。また、環境クズネット曲線に関わる大気質改善への寄与要因として、経済成長それ自体よりも環境管理能力が直接的であるという報告は、経済成長と環境改善の関係をとらえる上で重要な示唆と考えられる。

「企業経営と地域経済環境資産形成へのアプローチ評価(企業経営の視点から)」(築島武彦、梅内拓生)では、企業における従来の収支決算という財務会計的枠組みに、生涯コスト、心の資産、人材育成費用、環境的便益、社会連帯価値等の外部経済効果を取り入れ、地域社会への貢献度をより明確な形で位置づける「地域経済環境資産」の基本的理念について報告された。今後重要となる企業の社会的責任あるいは環境会計の考え方を推し進める評価アプローチの一つとしてとらえることができるであろう。

「日本における環境教育の評価の現状と課題：ESDの評価を視野に入れて」(藪並郁子、阿部治)では、環境教育関連の研究論文・発表内容を分析することで、現在の日本における環境教育の評価理念・手法を目的、対象及び指標の観点から整理するとともに、「持続可能な開発のための教育(ESD)」に向けた評価の課題や方向性が提言された。特に、地域の文脈に合った評価指標の設定、そしてプロセス評価とインパクト評価の適切なバランスの確保の必要性が強調された。

自由論題セッション 「教育評価」

座長 橋本 昭彦（国立教育政策研究所）

本セッションは3件の報告の後、質疑応答があった。

第一報告「義務教育課程における学校評価導入に関する考察」（小久保純一会員）は、民間企業の経営管理のノウハウや、学校経営・学校評価の諸研究に学びつつ、「マネジメント思考に基づく学校評価システムの構築」を提起した。学校評価によって「開かれたシステムへの転換」と「コミュニケーションの活発な組織づくり」を実現することを訴えた。

第二報告「学校評価における第三者評価の意義とその課題」（森田祐司・武久顕也・村上恵美・植田みどり会員）は、学校評価の政策的動向を整理し、教委等への調査によって得られた学校評価導入の現状や問題点を紹介した上で、第三者評価導入の課題を、その「目的」「品質」「費用」「担い手」や、戦略的な在り方などに分けて提示した。

第三報告「モンゴル農村部における教育の質的向上：国連・ユネスコプロジェクト評価結果からの政策提言」（山口しのぶ・矢野智子会員）は、2年間に及ぶ「冷害を受けたモンゴル国3県における学校修復及び学校長・教員のための再研修」プロジェクトの評価結果を、現地の風土に適する教育政策提言に活かすという、発展的な報告であった。

それぞれにつき若干の質問があり、それに対して報告の補充が行われた。質疑の正確な記録が用意できないので、個人的な所感を述べる。第一報告は、先行研究やデータ解釈に粗さがみられたが、鮮明な問題意識を活かすためには、報告者が公募教育長として経験されたことの客観的な報告にいつそうのウェイトを置き、そこから独自の知見を構築することを望みたかった。第二報告は、共同プロジェクトとして系統的に準備され、政策から地方教委・学校の実情までを視野に入れた労作である。現在国が進めている公立学校への「第三者評価」の方向性づくりから、時宜を得た報告であった。第三報告も、国際協力事業の紹介に始まって、その評価結果やそこから導き出せる政策提言を披露した濃厚な報告であった。教育の政策評価の実践的な事例研究としての優れた典型といえるが、中身が豊富すぎて、質疑の焦点も容易に定まらなかった。座長の取りさばきの腕前を反映して、質疑が全体に不十分に終わったが、報告者・参会者各位のご協力に感謝したい。

自由論題セッション 「公共財政」

座長 林 薫（文教大学）

本セッションでは二つの報告が行われた。なお、当初三報告が予定されていたが、一報告は直前に辞退となった。

最初の報告は上野宏会員（南山大学）による「途上国の公共財政制度改革と財政執行の評価方法2：PEFA業績測定枠組みへの業績評価報告書の考察」である。これは2005年6月に発表された同評価測定枠組の1年半にわたる評価状況をテーマにしたもので、評価枠組みの概観を行った前回評価学会全国大会のフォローアップという位置づけができる。報告においてはタンザニア、ザンビア、アフガニスタン、モザンビーク、キルギスの5カ国に関するPEFAの評価報告書を分析対象とし、まだどの報告書も内容は相当異なり標準的な評価手法の形成段階にあること、31指標による評価はPFMの優先順位を決定するのに有用であること、評価提言のみならず政策提言が報告書には含まれており、報告書の価値を高めていること、などの分析結果の報告が行われた。質疑応答では、PFM評価の最新の動向についての関心が多く示された。

次の報告は山下道子会員（（財）国際開発センター）による「最貧国における『一般財政支援』の評価」で、これは英国国際開発庁（DfID）の提唱で開始された一般財政支援の評価コンソーシアムのイニシアテ

ィブでまとめられた専門家による一般財政支援報告書に関するものである。山下会員は同報告書概要を述べつつ、一般財政支援は万能薬ではないが、途上国の政府のオーナーシップと説明責任を強化したこと、一般財政支援が途上国の貧困層へのサービス改善に寄与していること、一般財政支援に係る技術協力の効果は疑わしいことなどの評価結果を紹介した。これに対して、質疑応答では同報告書が形成評価であることに留意すべきであることなどについての質問や議論が行われた。総じて、財政支援や財政管理に係る評価の最新の動向を紹介したという点において意義が大きなセッションであったと思われる。

特別セッション「評価インターンによる調査報告」(中間報告)

座長 源 由理子 (国際開発コンサルタント)

本セッションでは、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」のもと、各インターンによる調査報告が行われた。

まず、立教大学大学院の松尾武司氏により「市民フォーラム21・NPOセンター」の組織評価について発表があった。同組織の中期ビジョンを達成するための組織力強化や人材育成に関する提言策定と目的としたもので、事業担当者と顧客に対する聞き取り調査、役員に対するアンケート調査結果をベースにデータ比較を行い、その差異から運営上の課題や必要な人材コンピテンシーを明らかにする方法が提示された。次に、呉市立昭和東小学校に派遣された竹下雅美氏(学習院女子大学大学院)、山崎その氏(同志社大学大学院)から、「ロジックモデルによる学校評価の改善」について発表があった。同報告では、教員及び保護者、地域住民に対する事前アンケート調査と聞き取り調査、児童に対する聞き取り調査をベースに、学校経営計画、学校評価に対する問題点を明らかにし、ロジック・モデルをツールとしてわかりやすい学校経営計画書と評価表を提案する過程が報告された。最後の発表となった小平谷中学校の評価では、丸山真司氏、市野緑氏(以上、東京工業大学大学院)から「授業改善を目的とした生徒による授業評価システムの開発と実践」と題し、調査の途中経過について発表があった。現行の授業評価に関する問題点の整理を目的に教員アンケート調査を実施し、その因子分析結果に基づき教員の視点から捉えた授業評価の課題が報告された。今後は再設計された授業評価の実施とその評価を踏まえ授業評価のあり方を提言する予定である。これらの報告に対し、それぞれのコメンテーターである西出順郎会員(琉球大学)、長尾真文会員(広島大学)並びに高千穂安長会員(玉川大学)から、評価倫理、データ収集方法、データ分析手法やその結果の提示方法といった多角的な視点からアドバイスを得た。

以上三つの報告は対象、目的が異なるも共通点としては組織の改善に評価を活用したいというインターン受け入れ先の思いがベースにある。その思いに、第三者としての客観性を保ちながらどれだけ有効なデータを提示できるのかが問われるところであるが、3月の最終報告に向けて、今回の発表に対するコメントを反映したより質の高い評価調査を期待したい。

企画委員会

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

1. 研究受託事業

前年度に引き続き、2006年度も国際協力機構より「事業評価の二次評価」業務を受託した。国際協力機構はすべての技術プロジェクト案件について終了時評価を行っているが、年間約50件とその数が多いところから、内部評価として実施している。しかし、評価内容の透明性、客観性を担保するために、作成された評価報告書を外部有識者事業評価委員会で確認する事になっている。外部有識者事業評価委員会の下に作業部会を設け、計量的な分析を加味して、公正なレビューを行っているが、この部会の作業を学会が請け負ったものである。

この成果の一部は2006年12月開催の第7回全国大会で「二次評価における評価視点の分析」(自由論題セッション 評価手法1)として一部発表された。また、この結果は2006年度末に公表された国際協力機構「事業評価年次報告書2006」の中に収録されている。

2006年度はさらに、外務省から、「無償資金協力におけるプロジェクト・レベル事後評価に係る評価調査(外務省による評価の二次評価)」業務を受託した。外務省では2005年度から無償資金協力プロジェクトの事後評価を開始したが、各大使館職員が評価に当たるところから、評価内容の透明性、客観性を担保するために、第三者に評価結果の二次評価を委託したものである。この二次評価結果は、2006年度末に公表されている。なおこの業務は一般競争入札に参加して受託したものである。

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が大きいの。今後とも、このような受託事業を積極的に拡大していきたいと考えている。

なお、学会として事業を受託することが適切であるかどうかの判断基準、及び受託事業の品質確保については学会が定めた「受託評価関連事業実施要領」に基づいている。

2. 顕彰事業

学会賞には論文賞、奨励賞、功績賞の3種がある。

2006年度は学会賞受賞者選考細則第6条により、2006年8月22日付けで学会ホームページ上に学会賞候補者の公募を行い、10月31日まで自薦、他薦を受け付けた。11月6日に学会賞審査委員会を開催した。論文賞、奨励賞、功績賞のそれぞれについて、申し合わせに基づき、功績、論文、発表原稿等公表された資料、アンケート結果、及び、日頃の学会活動などを総合的に検討し、さらに、受賞候補者の意思確認のためメールにて議論を継続し、11月15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6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、田中弥生会員を選考した。

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6回全国大会、及び春季第3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、受賞候補者として源由理子会員を選考した。

功績賞は、前年度までの学会活動への貢献に基づき、推薦されるものとする、と定められているが、学会賞選考内規に沿って、廣野良吉会員、長尾真文会員を選考した。

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4人の受賞者には、12月2日の総会場で、学会から賞状及び副賞が授与された。

企画委員会委員長 牟田 博光

研修委員会

活動方針

研修委員会では、「評価の普及および評価に関する人材育成に寄与すること」を目的として、各種セミナーの開催、研修プログラムの開発、教材の作成、講師の派遣、その他人材育成に関わる活動の実施支援を行うことを基本方針としている。2006年度は以下のような活動を展開した。

活動状況

1. 「評価インターン出前サービス」の企画・実施

「評価インターン出前サービス」は、学会の学生会員（大学院生）に対して「実際に評価を経験する」機会を提供すべく、「評価を外部に委託したい」と考えている地方自治体やNGO・NPO、学校等で実際に評価を体験させていただく事業である。今年度は、9月より各インターンによる実質的な調査がスタートした。なお、インターンによる調査結果は2006年12月の全国大会（東洋大学）の場で報告され、2007年3月には、最終報告書がとりまとめられた。

(1) 市民フォーラム21

「市民フォーラム21・NPOセンターの組織評価」

松尾武司（立教大学）

(2) 呉市立昭和東小学校

「ロジックモデルによる学校評価の改善」

竹下雅美（学習院女子大学）、山崎その（同志社大学）

(3) 東京都小平市立小平第二中学校

「授業改善を目的とした生徒による授業評価システムの開発と実践」

丸山真司、市野緑（以上、東京工業大学）

2. 学校評価研修プログラムにかかる認証制度説明会の実施

本学会が実施している学校評価研修プログラムの認証制度の説明会を都道府県及び政令指定都市及び中核市の教育委員会関係者を対象に以下のとおり実施した。

(1) 日 時：2007年1月19日（金）13：00～15：00

(2) 場 所：経団連会館 9階906号室

(3) 講 師：岸本哲哉氏（文部科学省初等中等教育局学校評価室長）

長尾眞文（広島大学、研修委員会委員）

石角剛氏（広島県立教育センター）

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」ことを目的とします。評価学会の活動のみならず広く評価に関わる情報も会員の皆さんと共有すべくあらゆる機会を捉えてその実現を目指します。

2. 活動状況

(1)「学会活動の強化と学会の健全な運営を目指して会員獲得に努めます」

会員獲得のための学会説明などのツールについては、ホームページを第一に考えていますが、国、自治体、企業などの説明ツールとしては、マス・メディアを使ったパブリシティ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」

評価を行う各機関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、評価資料の発表後、内容について転送する機関を増やしていきます。

(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、2006年6月24日理事会において計画推進について了解を得ました。

(4)「ホームページの有効性を評価します」

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5回委員会を2006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、効率性を高めていきます。

広報委員会委員長 高千穂 安長

国際交流委員会

1. 2006年10月18日より20日まで、2006 Asian Regional Forum on Aid Effectivenessがマニラで開催された。本地域フォーラムはアジア開発銀行、日本政府、DIFD、世界銀行の共催で、OECD・DACのパリ宣言を受けて、援助の有効性を援助側、被援助側、NGO、コミュニティーグループ等で議論するものであった。分科会のODA評価セッションでは、廣野良吉国際交流委員会委員長が議長を務め、三輪徳子氏、湊直信氏がプレゼンテーションを行った。廣野議長からアジア・パシフィックの評価学会のネットワークとしてのAPEA (Asia Pacific Evaluation Association) 設立の提案がなされ了承された。
2. 国際交流委員会では上記のAPEA設立に関して、2回のブレインストーミングを踏まえて、学会全体に賛同者を募ることとし、3月2日に関心を持つ会員が集まり会合を持った。今後、APEA設立に向けて積極的に活動を行っていく予定である。なお、APEA設立への委員会活動は、本学会設立以来活動方針として会員総会で承認されている「国際開発評価学会への積極的対応」の一環である。

国際交流委員会委員長 廣野 良吉

日本評価学会春季第4回全国大会のご案内

第4回全国大会

実行委員会委員長 和田 義郎

プログラム委員長 牟田 博光

来たる2007年6月2日（土）、日本評価学会春季第4回全国大会が学術総合センター（東京都千代田区）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： 2007年6月2日（土）
2. 場所： 学術総合センター（東京都千代田区一ツ橋2-1-2）

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品川区に置く。

(目的)

第3条 本法人は、広く国民一般に対する透明性確保、説明責任遂行のため、公共的な活動やNPO等の社会的な活動に関する評価の重要性が強く認識されていることに鑑み、評価に関する研究及び応用を促進し、評価に関する関連機関等との情報交換を図るとともに、この分野の学問の進歩発展及び評価に携わる人材の育成等に貢献することにより、広く一般市民に対し評価活動の向上と評価の普及に寄与することを目的とする。

(特定非営利活動の種類)

第4条 本法人は、前条の目的を達成するため、次の種類の特定非営利活動を行う。

- (1) 社会教育の推進を図る活動
- (2) 学術、文化、芸術又はスポーツの振興を図る活動
- (3) 国際協力の活動
- (4) 男女共同参画社会の形成の促進を図る活動
- (5) 経済活動の活性化を図る活動

(事業の種類)

第5条 本法人は、第3条の目的を達成するため、特定非営利活動に係る事業として次の事業を行う。

- (1) 評価に関する調査・研究事業
広く一般市民の参加を得て客観性を保持し得る評価手法、評価基準の開発を行う。
- (2) 評価に関する研修事業
評価に関する学術研究会、講演会、国際シンポジウム、研修などを開催し、評価人材の育成に努める。
- (3) 評価に関する交流事業
評価研究者、評価実務者、評価利用者等評価に関連する内外の人々との交流、情報交換の場として国際交流や全国大会を実施する。
- (4) 評価に関する広報・普及事業
評価に関する情報を広く知らしむべく、ホームページの作成、学会誌その他刊行物の発行等、普及啓蒙活動を行う。
- (5) 評価実施事業
公平、中立の観点に立ち、法人として第三者評価を実施する。
- (6) 評価に関する顕彰事業
評価に関する研究および評価活動の普及に貢献した人を広く顕彰する。
- (7) その他、本法人の目的を達成するために必要な事業

第2章 会員

(種別)

- 第6条 本法人の会員の種別は、次の5種とし、正会員をもって特定非営利活動促進法（以下「法」という。）上の社員とする。
- (1) 正会員 本法人の目的に賛同して入会した個人。
 - (2) 学生会員 本法人の目的に賛同して入会した、原則として大学以上の学生で、学生会員を希望する者。尚学生会員は、卒業時に本人が希望する場合は会長に申し込みをおこなったうえで、正会員となることができる。
 - (3) 賛助会員 本法人の目的に賛同し、その事業を後援する団体。
 - (4) 名誉会員 本法人に功労のあった者及び広く評価分野に関連ある分野における学識経験者で理事会の推薦に基づき総会の承認を経た個人。
 - (5) 准会員 本法人の目的に賛同して入会した個人で、最大3年間を期限に本法人の活動への参加を希望する者。3年以上の継続を本人が希望する場合は会長に申し込みをおこなったうえで、正会員となることができる。

(入会)

- 第7条 会員の入会について、特に条件は定めない。
- 2 正会員、学生会員、賛助会員、准会員として入会しようとする者は、会長が別に定める入会申込書により会長に申し込むものとする。また、名誉会員として推薦された者は、入会の手続きを要せず、本人の承諾をもって会員となるものとする。
 - 3 理事会は、前項の申し込みがあったとき、正当な理由がない限り、入会を認めなければならない。
 - 4 会長は、第2項のものの入会を認めないときは、速やかに理由を付した書面をもって本人にその旨を通知しなければならない。

(会費)

- 第8条 正会員、学生会員、賛助会員、准会員は、総会において別に定める会費を納入しなければならない。

(会員の資格の喪失)

- 第9条 会員が次の各号の一に該当する場合には、その資格を喪失する。
- (1) 退会届の提出をしたとき。
 - (2) 本人が死亡し、若しくは失踪宣告を受け、又は会員である団体が消滅したとき。
 - (3) 継続して2年以上会費を滞納したとき。
 - (4) 除名されたとき。

(退会)

- 第10条 会員は、会長が別に定める退会届を会長に提出して、任意に退会することができる。

(除名)

- 第11条 会員が次の各号の一に該当する場合には、総会の議決により、これを除名することができる。
- (1) この定款に違反したとき。
 - (2) 本法人の名誉を毀損し又は設立の趣旨に反する行為を行なったことにより、総会において、

除名すべきものと認められた場合

- 2 前項の規定により会員を除名しようとする場合は、議決の前に当該会員に弁明の機会を与えなければならない。

(抛出品品の不返還)

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 役員の任期は、2年とする。ただし再任を妨げない。

- 2 補欠又は増員により就任した役員の任期は、前項本文の規定にかかわらず、前任者又は現任者

の残存期間とする。

3 役員は、辞任または任期満了後においても、後任者が就任するまでは、その職務を行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) 役員を選任又は監事の解任、役員職務
- (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ができない。

(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) 会長が必要と認めたとき。
- (2) 理事総数の2分の1以上から理事会の目的である事項を記載した書面により招集の請求があったとき。

3 理事会は、理事現在数の3分の1以上の出席がなければ、開会することができない。

(理事会の招集)

第34条 理事会は、会長が招集する。

2 会長は、前条第2項第2号の場合にはその日から20日以内に理事会を招集しなければならない。

3 会長は、緊急に理事会を招集する必要がある場合において、やむを得ない事情によりこれを開催できないときには、理事の承諾を得て、書面により議決を得ることができる。この場合、理事会は開催されたものとみなす。

4 理事会を招集するときは、会議の日時、場所、目的及び審議事項を記載した書面により、開催の日の少なくとも5日前までに通知しなければならない。

(理事会の議長)

第35条 理事会の議長は、会長がこれにあたる。

(理事会の議決)

第36条 理事会における議決事項は第34条第4項の規定によってあらかじめ通知した事項とする。

2 理事会の議決は、出席理事の過半数の同意でこれを決し、可否同数のときは、議長の決するところによる。

(理事会の表決権等)

第 3 7 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 3 8 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 3 9 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 4 0 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 4 1 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 4 2 条 この法人の会計は、特定非営利活動に係る事業会計とする。

(事業年度)

第43条 本法人の事業年度は、毎年10月1日に始まり翌年9月30日に終わる。

(事業計画及び予算)

第44条 本法人の事業計画書及び収支予算書は、会長が作成し、理事会の議決を経た後、毎事業年度の開始前に総会の議決を得なければならない。ただし、やむを得ない事情により、当該事業年度開始前に総会を開催できない場合にあっては、理事会の議決によることを妨げない。この場合、当該事業年度の開始の日から90日以内に総会の議決を得るものとする。

2 会長は、前項の事業計画書及び収支予算書を変更しようとするときは、理事会の議決を得なければならない。

(暫定予算)

第45条 第43条の規定にかかわらず、やむを得ない理由により予算が成立しないときは、会長は、理事会の議決を経て、予算成立の日まで前事業年度の予算に準じ収入支出することができる。

2 前項の収入支出は、新たに成立した予算の収入支出とみなす。

(予備費)

第46条 予算超過又は予算外の支出に充てるため、予算中に予備費を設けることができる。

2 予備費を使用するときは、理事会の議決を経なければならない。

(事業報告及び決算)

第47条 本法人の事業報告書、財産目録、貸借対照表及び収支決算書等決算に関する書類は、会長が毎事業年度終了後遅滞なくこれを作成し、監事の監査を受け、理事会の議決を経た後、当該事業年度終了後90日以内に総会の決議を得なければならない。

2 決算上剰余金を生じたときは、次事業年度に繰り越すものとする。

(臨機の措置)

第48条 予算をもって定めるもののほか、借入金の借入れその他新たな義務の負担をし、又は権利の放棄をしようとするときは、理事会の議決を経なければならない。

第7章 定款の変更、解散及び合併**(定款の変更)**

第49条 本法人が定款を変更しようとするときは、総会に出席した正会員の4分の3以上の多数による議決を経、かつ、法第25条第3項に規定する軽微な事項を除いて所轄庁の認証を得なければならない。

(解散)

第50条 この法人は、次に掲げる事由により解散する。

- (1) 総会の決議
- (2) 目的とする特定非営利活動に係る事業の成功の不能
- (3) 正会員の欠亡

- (4) 合併
- (5) 破産
- (6) 所轄庁による設立の認証の取消し
- 2 前項第1号の事由により本法人が解散するときは、正会員総数の4分の3以上の承諾を得なければならない。
- 3 第1項第2号の事由により解散するときは、所轄庁の認定を得なければならない。

(残余財産の帰属)

第 5 1 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 5 2 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 告 告

(公告の方法)

第 5 3 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 5 4 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 1 0 章 雑 則

(細則)

第 5 5 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 5 6 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 5 7 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の投稿についても、常任編集委員会による掲載の判断は可・不可の二者択一で行うこととする。但し、場合によっては編集委員会の判断で、小規模の修正による掲載も認める。「研究論文」としての掲載が適当でないと判断された場合でも、投稿者が希望すれば、常任編集委員会は「研究ノート」あるいは「実践・調査報告」としての掲載を決定できる。

(投稿要領の作成公表)

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

(配布先)

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

(抜刷の配付)

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

(インターネット上の公開)

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

(著作権)

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

(事務局)

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

(以上)

『日本評価研究』投稿規定

2003.4.18改訂

2002.3.25改訂

2001.9.9改訂

1. 『日本評価研究』(The Japanese Journal of Evaluation Studies)は、評価に関する論文、論考、調査報告等を掲載する。
2. 『日本評価研究』は、会員間の研究成果交流の場を提供し、内外における評価研究の一層の発展に資することを主目的として発行されており、原則として会員による寄稿を掲載する。なお、依頼原稿を除き、ファーストオーサーは学会員でなければならない。
3. 投稿された原稿は、編集委員会の責任において審査を行ない、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

〒140-0002 東京都品川区東品川4 - 12 - 6 日立ソフトタワーB 22F
財団法人 国際開発センター内 日本評価学会事務局
TEL: 03-6718-5931
FAX: 03-6718-1651
E-mail: jes_info@idcj.or.jp

『日本評価研究』執筆要領

2002.9.18改訂

2002.3.25改訂

1. 本文、図表、注記、参考文献等

(1) 論文等の記載は次の順序とする。

日本語原稿の場合

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

第2ページ以下：本文、謝辞あるいは付記、注記、参考文献

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

英文原稿の場合

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

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

(2) 本文の区分は以下のようにする。

例1（日本語）

1 .

(1)

(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

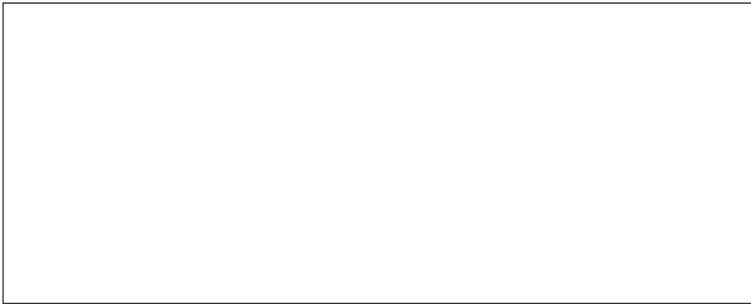
図1 州における生徒数の推移



(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ない

論文の構成、表現（用語、引用、図表等）が適切でない（または整合性がとれていない）

6. 判定

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

(1) 掲載可とする。

(2) 小規模の修正による掲載可とする。

(3) 大幅な修正による掲載可とする。

但し、(総説/研究論文/研究ノート/実践・調査報告)として掲載を考慮できる。

(4) 掲載不可とする。

但し、(総説/研究論文/研究ノート/実践・調査報告)として掲載を考慮できる。

英文学会誌の刊行について

2007.1.20
『日本評価研究』
編集委員会

本学会では、学会設立当初以来、学会誌を年2回刊行し、恒常的に学会員の評価に関する研究及び実践活動の成果を国内外に学会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努めてきました。しかし、学会誌に掲載した論文等は和文のものが大部分であり、より一層の国内外での評価に関する知識、知見の共有化のためには、英文誌を下記により刊行することとしました。会員皆様の投稿をお待ちいたします。

記

1. 現在の学会誌の刊行を、和文誌と英文誌に分け、年度において和文誌2回、英文誌1回、刊行する。
2. 刊行時期は、和文誌は、従来どおり、3月および9月、英文誌は、12月を目処に行う。英文誌創刊号の原稿締め切りは、平成19年6月30日とする。
3. 学会員の評価に関する研究および実践活動の成果を国内外に広く発表するために、和文誌に掲載された論文等を、英文に翻訳し、英文誌に掲載すること、また、英文誌に掲載された論文等を、和文に翻訳し、和文誌に掲載することを認める。但し、翻訳した論文等は、翻訳言語での体裁等を満たしていることを条件とする。また、翻訳は、執筆者が各自の責任において行なうものとし、和文誌または英文誌に掲載された論文等の翻訳であることを掲載誌に明記することとする。
4. 投稿および原稿依頼等は、別添の刊行規定等によるものとする。
 - (1) Publication Policy of the Japanese Journal of Evaluation Studies
 - (2) Information for Contributors (For English papers)
 - (3) Writing Manual of the Japanese Journal of Evaluation Studies (For English papers)
 - (4) Referee-Reading Guideline

以上

別添 (1)

Publication Policy of the Japanese Journal of Evaluation Studies

Last revised on 15th February 2005

(The Purpose and the Name)

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

(Editorial Board)

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

(Editorial Policy)

- 8 . The evaluation study, as a principle, is published twice a year.
- 9 . The evaluation study is printed on B5 paper, and either in Japanese or English.
- 10 . Papers published in the evaluation study are categorized as five types;
 - (1) Review
 - (2) Article
 - (3) Study Note
 - (4) Report
 - (5) Others
- 11 . The qualified contributors are members of the evaluation society (hereinafter referred to as “ members ”) and persons whose contribution is requested by the standing editors. Joint submission of members and joint submission of non-members with a member as the first author are accepted. Submission by the editors is accepted.
- 12 . Submitted manuscripts are treated as the above categories, however, the standing editors will decide based on the application of the contributors and the following guidelines;
 - (1) “ Review ” is a paper, which provides an overview of evaluation theory or practice. The editorial board will make the decision regarding publication.
 - (2) “ Article ” is considered as a significant academic contribution to the theoretical development of evaluation or understanding of evaluation practice. The standing editors committee makes adoption judgments following the referee-reading process described in the next section.

- (3) “ Study note ” is a discussion equivalent to the intermediate outputs of a theoretical or empirical enquiry. The standing editors committee makes adoption judgments following the referee-reading process described in the next section.
- (4) “ Report ” is the study report related to a practical evaluation project or evaluation. The standing editors committee makes adoption judgments following the referee-reading process described in the next section.
- (5) “ Others ” includes requested papers for special editions organized by the editorial board and announcements from editorial board to members regarding publication.
- 13 . The editorial board selects two referee readers. For the “ article ”, the editorial board makes adoption judgments referring to the results from referee readings and comments provided by one editor assigned by the editorial board. For “ review ”, “ study note ”, “ report ” and “ others ”, the editorial board makes adoption judgments referring to the results from referee readings.
- 14 . When editors submit a manuscript, the editors are not allowed to attend any of the standing editors committee meetings or editorial board meetings regarding the manuscript.
- 15 . The standing editors have alternative of approval or not-approval for adoption judgment of manuscripts submitted to any categories. However exception is permitted if the editorial board approves the publication after minor rewrite. Even if the manuscripts are considered insufficient as an “ article ”, standing editors can decide whether the manuscripts are published as a “ study note ” or “ report ” if the authors wish to publish.

(Formulation and Release of Submission Procedure)

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

(Distribution)

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

(Distribution of the Printed Manuscript)

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

(Release on the Internet)

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

(Copyright)

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

(Office)

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

別添 (2)

Information for Contributors (For English papers)

Last revised on 18th April 2003

- 1 . “ The Japanese Journal of Evaluation Studies ” is the publication for reviews, articles, study notes, and reports relating to evaluation.
- 2 . “ The Japanese Journal of Evaluation Studies ” is primary published to provide opportunities for members of the Japan Evaluation Society (hereinafter referred to as “ members ”) to exchange findings, and to contribute to further development of the study of evaluation both domestically and internationally. As a principle, this journal publishes the contributions submitted by the members. With the exception of requested papers, the first author must be a member.
- 3 . Adoption judgments of the manuscript are made at the discretion of the editorial board. Comments from two referee readers who are appointed for every manuscript are referred to in the screening process (the editorial board requests referee readers without notifying the author of manuscript).
- 4 . Payment for the manuscript is not provided.
- 5 . Papers published in “ The Japanese Journal of Evaluation Studies ” are released on the Internet at homepage of this academic society.
- 6 . Regarding submission, manuscripts must be identified as one of the following categories: 1) article, 2) review, 3) study note, 4) report, and 5) others. However, the final decision of the category is made by the editorial board.
 - “ Article ” is considered as a significant academic contribution to the theoretical development of evaluation or understanding of evaluation practice.
 - “ Review ” is a paper which provides an overview of evaluation theory or practice.
 - “ Study note ” is a discussion equivalent to the intermediate outputs of a theoretical or empirical study in the process of producing an “ article ”.
 - “ Report ” is the study report related to a practical evaluation project or evaluation.
 - “ Others ” are manuscripts for special editions requested by the editing committee.
- 7 . Manuscript Submission
 - (1) Manuscripts may be written in either Japanese or English.
 - (2) Correction by the author is only for the first correction.
 - (3) English manuscripts should be submitted only after the English has been checked by a native speaker.
 - (4) Submit four hard copies (A4 size) of the manuscript. Contact information including mailing address, telephone number, fax number, and e-mail address, and the category of the manuscript should be clearly stated.

For approved manuscripts, after necessary rewriting, the author needs to submit two hard copies of the final paper as well as a text file saved on a DOS/V formatted floppy disk. Original figures, charts, and maps should be provided.
 - (5) Total printed pages should not exceed 14 pages. Any cost incurred by printing more than 14 pages must be covered by the author.
 - (6) The layout for English papers should be 30 mm of margin at left and right side, 10pt for font size, 43 lines on A4 paper (about 500 words per page). An abstract of 150 words should be attached to the

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

8 . Mailing address

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

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

E-mail: jes_info@idcj.or.jp

別添 (3)

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

Revised on 18th September 2002

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

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

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

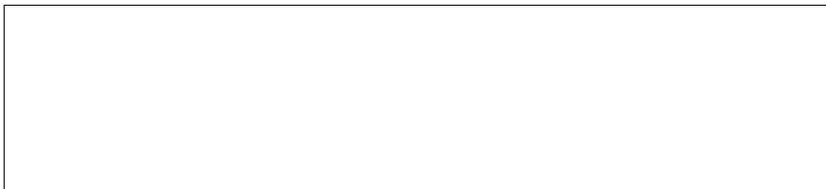
Second page: The main text; acknowledgement; notes; references

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

- 1.
- 1.1
- 1.1.1
- 1.1.2

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

Figure 1 Number of Students in the State of



Note:

Source:

Table 1 Number of Accidents in the State of

Note:

Source:

(4) Citation of literature in the text should be, (Abe 1995, p.36) or (Abe 1995).

(5) Note in the text should be, (-----.¹)

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

Note

1 -----.

2 -----.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

別添 (4)

Referee-Reading Guideline

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

1. Content of the Referee-Reading Guideline

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

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

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

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

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

3. Items of Consideration in Referee-Reading

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

- (1) Importance and utility of the theme
 - (2) Originality of the study
 - (3) Structure of the logic
 - (4) Validity of verification and methodology
 - (5) Contribution to evaluation theory and practice
- For the article, all of above five are considered.
 - For the study note, especially (1), (2), (3), and (4) are considered.
 - For the report, especially (1), (3), and (5) are considered.
 - For the review, especially (3) and (5) are considered.

4. Attentions in submission of manuscript

Besides above five viewpoints, basic completeness as a paper is also considered, for example;

- appearance of the paper is organized
- written according to the writing manual
- described simply and distinctive
- verification data is appropriately used
- notes and references are corresponding with the text
- terminology is appropriately used
- no wording and grammatical mistakes
- no errors and omission
- no punctuation mistakes
- expression in English abstract is appropriate
- word count is according to the manual

The above mentioned forms and contents are also considered. There have been cases in which graduate students and practitioners posted without organizing the manuscripts as a paper. On those occasions, referee-reading was not conducted. Necessary consultation is strongly recommended prior to submission.

5. Judgment Cases in Referee-Reading

(1) In the case of the manuscript which is considered acceptable for the publication but is not yet complete:

The referee reader should evaluate carefully whether the paper can contribute to the development of evaluation theory or evaluation studies.

- Verification is lacking but the theory and formulation are useful for academic development.
- Analysis lacking but useful for formation and promotion of new theory.
- The literature review is not of a high standard but, the overall study is meaningful.
- Comparative study is not up to standard but is meaningful as an example of application.
- Analysis is lacking but it is meaningful as an evaluation of socially and historically important cases.
- Analysis is lacking but it is meaningful as an evaluation of particular social activities.
- Organization and expression are not up to standard as a paper but the contents are worthy to evaluate.
- Logic is not strong enough but useful in practice.
- The paper has significance as a report.

(2) In case of the manuscript which is considered as difficult for publication:

- Awareness of the issue or setting of the problem is indecisive.
- Understanding or analytical framework of notion of basic terminology is indecisive or inappropriate.
- There is a lack in credibility of data for the grounds of an argument.
- There is no clear point of an argument or appropriateness of proof.
- Organization of the paper and presentation (terminology, citation, chart, etc) are inappropriate (or not consistent).

6. Judgment

The final decision will be made on publication at the standing editors committee following one of four patterns (listed below). However, these judgments are not based on the number of errors but on the strength of the overall report. In the case of (3) and (4), there is a possibility to be published as a different type of paper. If it is published as a different type of paper, major rewrite concerning the number of words may be required.

- (1) The paper will be published.
- (2) The paper will be published with minor rewrite.
- (3) The paper will be published with major rewrite, however as a different type of paper (review, article, study note, or report).
- (4) The paper will not be published; however there is the possibility that it will be published as a different type of paper (review, article, study note, or report).

～ 投稿案内 ～

日本評価学会では、「日本評価研究」掲載のための投稿原稿を募集しております。1年に2回（3月末／9月末）の投稿締切を設けておりますので、ご興味をお持ちの方は投稿規定・執筆要領をご参照のうえ、奮ってご投稿ください。

投稿先：特定非営利活動法人日本評価学会 事務局
〒140-0002 東京都品川区東品川4-12-6 日立ソフトタワーB 22F
財団法人国際開発センター内
TEL: 03-6718-5931 FAX: 03-6718-1651

『日本評価研究』第7巻第1号

2007年3月30日

編集・発行 特定非営利活動法人 日本評価学会
〒140-0002 東京都品川区東品川4-12-6
日立ソフトタワーB 22F
財団法人国際開発センター内
TEL: 03-6718-5931 FAX: 03-6718-1651

印 刷 株式会社 研恒社

© 日本評価学会

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

The Japanese Journal of Evaluation Studies

Vol. 7, No. 1, March 2007

CONTENTS

Special Issue: Educational Evaluation

Editor's Note Hiromitsu Muta

Theoretical and Practical Issues in School Evaluation Masafumi Nagao

A Report on School Improvement Practice through Self-evaluation
Kengou Ishida, Etsuko Taira, Shinobu Sumimoto, Masafumi Nagao

For Further Improvement of Universities: Evaluation of Evaluation Program Implemented by
National Institution for Academic Degrees and University Evaluation in Japan
Takahiro Saito, Takayuki Hayashi

A Theoretical Comparative Analysis on Educational Evaluation in Developing Countries:
Perspectives of International Student Assessment, School Survey, and Household Survey
Mikiko Nishimura

Evaluation of Education Improvement Projects from the Viewpoints of Contribution to
Community Empowerment Yoko Ishida

Utilizing Participatory Evaluation in Non-formal Educational Cooperation Projects
– Relevance of Stakeholders Involvement in Evaluation Process – Yuriko Minamoto

An Evaluation Study of Program-Based Approaches (PBAs) for Development Assistance in
the Primary Education Sector: A Case Study of the Education for All -Fast Track Initiative
(EFA-FTI) in Honduras and Vietnam Satoko Miwa

Articles

Three-dimensional Analysis Incorporating the Weight of Factors for Evaluation
– A Case for the University Course Evaluation –
Atsuko Hoshino, Shun-ichi Kitahara, Kouji Shingyouchi, Kazuhisa Adachi,
Masayasu Watai, Hiromitsu Muta

Economic Burden of Disease and Cost Benefit Analysis of Emergency Medical Service
(EMS) System in Romania
Takao Ozaki, Marius Enescu, Moazzam Ali, Chushi Kuroiwa

A Causal Analysis for the Development Process of Social Capacity for Environmental
Management: The Case of Urban Air Quality Management
Kazuma Murakami, Shunji Matsuoka, Tatsuo Kimbara

A Study on Activities to Ensure Objectivity of Evaluation in Japan Tetsuya Nishimoto

Research Notes

Trends in Air Pollutant Concentrations and Related Environmental Legislation in India
Miyuki Nagashima

Diversification of Approaches to the Self-Evaluation of Undergraduate Education
Takeshi Kushimoto