

評価の理論と実践におけるプログラムの概念

政策評価とODA評価をめぐって

山谷 清志

あらまし

わが国の政策評価はアメリカで使われる「業績測定 'performance measurement' 」と(ABCや5段階の)「評定(rating)」を中心に普及している。しかし、他方でこれもアメリカで一般化した「プログラム評価」については、国土交通省や総務省の総合性確保評価などの少数の例外を除き、普及していない。また、ODA評価においては実施機関である国際協力機構や国際協力銀行がプロジェクト評価、外務省が政策レベル評価、プログラム・レベル評価を行うという分業体制を構築し、外務省はプログラム・レベル評価に力を入れることになっているが、これもまた思うように進んでいない。

わが国の評価の理論と実務でプログラム概念が普及せず、これを使った「プログラム評価」をうまく活用できていない理由のひとつは、プログラム概念それ自体の理解不足と混乱にあると本稿では考えている。第二次世界大戦後、世界銀行の融資を受けたことをきっかけに急速に普及し、日本語として違和感のないプロジェクトの概念とは対照的な状況にある。ここでは、こうした状況を明らかにしながら、本来あるべきプログラム評価のイメージを模索する。

1. はじめに

「プログラム」という言葉は最近になってようやく日本語になじんできた概念である。たとえば小学校の運動会のプログラム、ピアノ演奏会のプログラム、コンピュータのプログラムというように使われている。また、政治・行政の分野

でも、たとえば大阪府の教育改革プログラム(1999年4月)、林野庁の林政改革プログラム(2000年12月)、鹿児島県財政改革プログラム(2001年12月)、内閣府の金融再生プログラム(2002年10月)などがある。三重県(総合企画局政策推進室)では、知事の選挙公約であるマニフェストを「重点プログラム」という形でまとめ、その2004年から06年までの3年間の成果に対する外部評価を考えている。そして、小泉内閣では構造改革「プログラム」として7つ、すなわち民営化・規制改革、チャレンジャー支援、保険機能強化、知的資産倍増、生活維新、地方自立・活性化、財政改革を掲げた。

しかし、われわれが日本語でプログラムという言葉を用いるとき、実はそこに共通の理解があるわけではない。使う人によって、また背景の違いに応じて微妙な差異がある。あるときは「番組」や「予定(予定表)」という意味で使い、別のときには「計画」として使い、さらに操作や作業の「手順」「工程表」「マニュアル」というような意味でも使われる。また、中央府省でも地方自治体でも政策評価システムを導入する際に用いる政策体系図では、プログラムを「施策」と言い換えて使っている。ところが施策が政策と同じ意味で使われることがあり、また事業と施策の区別が付いていないことも多く、場合によっては施策と政策が同じ法律の条文で文言上両者の区別ができないときもある(たとえば男女共同参画社会基本法第5条)。

そしてこの混乱状況が、政策評価の深化、洗練を求めるとき、問題になってくる。抽象的な政策のレベルを対象にしたとき、評価は政治イデオロギーの対立や価値判断になりかねず、事業やプロジェクトの評価では「木を見て森を見ず」の

状態に陥るため、よりオペレーショナルな概念であるプログラムが目される。しかし、そのプログラムの評価とはいったい何を、どのように評価するのであろうか。

2. 「プログラム」概念への注目

そもそもここで議論しようとする「プログラム」概念は、20世紀末、わが国に現れたふたつの改革を背景として注目された。ひとつはODAの評価のあり方を再検討しようとする試みであり、もうひとつは2001年6月22日に参議院を全会一致で通過して成立した「行政機関が行う政策の評価に関する法律」に代表される政策評価の動きである。

2.1 ODAにおける「プログラム」

ODAの評価それ自体をめぐる議論は後に示す「略年表」に見られるとおり、1981年というかなり早くから議論されてきた。その後、評価に対するさまざまなニーズを反映して、ODA評価は一方の国別評価・援助実施体制評価・特定テーマ評価というように評価の対象別に分ける類型化と、他方の合同評価・有識者による評価・国際専門家による評価・被援助国関係者による評価・現地コンサルタントによる評価・シンクタンクによる評価・国際機関の評価・NGOとの共同評価・在外公館による評価というように評価主体による類型化とが混在し、それぞれの類型の中で複雑に展開していた。それは良く言えば精緻なメカニズム、悪くいえば素人に全く理解不可能なカオス状態であった。

こうした評価のあり方に対し、社会全般から「ODA評価はわかりにくい」「透明性と説明責任を高めるべき」という声が上がリ、ODA評価体制の見なおしの議論が、財政の問題や予算削減の問題といった比較的理解しやすい議論から、かなり難解な評価の方法論の問題に踏み込んで出てきた。それが2000年3月、経済協力局長の私的諮問機関である「援助評価検討部会」の下に置かれた「評価研究作業委員会」が提示したODA

評価改善に関する5課題である。この5課題とは 政策レベルの評価の導入とプログラム・レベルの評価の拡充、 評価のフィードバック体制の強化、 評価の人材育成と有効活用、 評価の一貫性の確保(事前から中間・事後に至る一貫した評価システムの確立)、 ODA関係省庁間の連携推進、である。

これら5課題を討議するために援助評価検討部会・評価研究作業委員会のもとに「ODA評価研究会」が設置された。研究者、公認会計士、援助評価の実務家からなるこの研究会は、とくにこの政策とプログラムのレベルにおける評価についてかなり多くの時間を割いて検討し、プログラムの評価という方法をもっと早く導入し、拡充する必要があるという意見について研究会参加者の認識が一致した。ただ、政策レベルやプログラム・レベルの評価については、その定義に一致が無く、曖昧なまま使用され、国際的に統一された具体的手法が確立していないことも判明した¹。

問題は、そもそも「プログラム評価」とはなにか、またそうした概念(認識ツール)がなぜ必要かということである。国際的な流れとしては、ODA評価は個々のプロジェクト評価に限らず、政策レベルやプログラム・レベルの評価も行うべきとの方向に向かいつつあり、現実にも多くのドナー国・国際機関は援助実施に際して、個々の具体的なプロジェクトの実施だけでなく、関連する複数のプロジェクトを有機的に組み合わせるアプローチ(プログラム・アプローチ)を取り入れるようになってきている。これに対応して、評価もプロジェクト・レベル評価('project evaluation'に対応)に加えて、プログラム・レベル評価を導入するようになってきている。さらに政策レベルの評価の重要性が国際機関の主催する会議、研究会で指摘され、援助機関全体として目指すべき方向性や戦略(たとえば貧困削減) 国別援助政策の評価が21世紀になって増えている。日本の外務省も、後ろに示した「資料: ODAの政策体系および評価対象」にあるように、政策レベル評価、プログラム・レベル評価、プロジェクト・レベル評価のそれぞれの考え方を整理しており、国際的な流れの延長線上にのっている。

¹ 援助評価検討部会・ODA評価研究会「ODA評価研究会報告書 - 我が国のODA評価体制の拡充に向けて - 」、2001年2月。

ただし、繰り返すが、現在のところ明確な政策レベルの評価、プログラム・レベルの評価の定義が存在するわけではない。実務において試行されているだけであった²。この状況を踏まえ、ODA 評価研究会ではその会議のはじめにプロジェクト、プログラム、ならびに政策レベルの評価など、用語の意味を再度整理統一する必要があるという提言を出した（第1回会合、2000年7月20日）。それに応えて外務省は2000年6月25日から7月17日にかけて、主要ドナー国（アメリカ・カナダ・スウェーデン・オランダ・デンマーク・フランス・ドイツ・イタリア・イギリス）と国際機関（世界銀行、国連開発計画、経済協力開発機構、国連食糧農業機関）に調査団を派遣し、政策レベル評価とプログラム・レベル評価の実態調査を行った。

ところで、ODA評価研究会の当初の議論では、もともと国際機関の場合プログラム評価と政策の評価の間にある境界線は曖昧であるため、ひとつひとつのプロジェクトの評価をつみあげてカントリー・ポートフォリオを作成し、それに基づいて対象国に対する援助政策の評価を行なう方法が実務上可能な考え方ではないかと見ていた。ただし、たとえばプログラム援助やノンプロジェクト援助、構造調整借款や商品借款、輸入資金や財政資金の不足を補う援助、あるいは発展途上国が必要とする商品の輸入代金や商品そのものの援助のように、評価手法が確定していない援助の評価は困難であることは理解していた。また、構造調整借款や商品借款の場合、たとえば被援助国の経済状態が良くなったとしても、それが援助によって改善されたのか、その援助以外の要因によるのかは必ずしも明らかにならないという指摘もあった。ちなみに現在はこうした技術的能力の限界を勘案して、プロジェクト評価を中心に、過去の援助実績の多い順から選択して評価を行っており、その選択基準は「フィードバックとして評価結果が役立つことであること」、あるいは「効果の出る時期が熟していること」である。あるいはまた、個々のプロ

ジェクトに焦点を当てて上位計画・目標との妥当性等を見る、言わばボトムアップ型の評価が「プロジェクト評価」であり、セクターあるいは地域の開発を目指すプログラムに焦点を当てて、その中で実施されている個々のプロジェクトとの関連性を評価するトップダウン型が「プログラム・レベル評価」であるという考えも、ODA 評価研究会において各委員から紹介されていた。

そして一応の結論として、ODA 評価研究会では次のような定義を試みた。政策決定レベルへのフィードバックを主目的とする評価が「政策レベルの評価」であり、個別プロジェクト群を包括して上位計画への貢献度を評価し、援助機関の援助方針・計画作成レベルへのフィードバックを主目的とするのが「プログラム・レベルの評価」であるという定義である。なおODA 評価研究会は、各委員の報告と外務省委託調査「政策レベル及びプログラム・レベルの評価に関する手法研究調査」の北米・欧州地域における調査結果、バングラデシュにおける試験的適用に関する調査報告を受けて議論・検討した結果、以下のような理論的に検討すべき課題を提示している。

- a .政策レベル、プログラム・レベル評価の早期導入、拡充が必要である。
- b .政策レベル、プログラム・レベル評価については、国際的に統一された具体的手法は未確立なので、わが国のODAに適した評価手法と体制の確立が重要課題である。
- c .援助政策、国別援助計画、プログラム、プロジェクト等の体系図を、事前段階から設定し、同時に指標設定やモニタリング方法を明確化する必要がある。
- d .評価に一貫性を持たせるため評価手法・手順に関するガイドライン、マニュアル、雛型等を整備すべきである。
- e .他のドナーや国際機関との協力による合同評価の利用と拡充を促すべきである。
- f .上位レベルの評価を効果的に行うため、援助国・被援助国双方の評価能力の向上と上

² ODA 評価研究会の提言がでた後、2003年3月に外務省経済協力局評価室が公表した『ODA 評価ガイドライン』によると、プログラム・レベル評価は「共通の目的を持った複数のプロジェクトなどの集合を対象とした評価であり、そこではセクター別評価とスキーム別評価がある」と定義されている。このガイドラインによればセクター別評価は、基本的に1カ国、1セクターにおけるODA活動の集合体（医療、保険、インフラ整備などのセクター）を対象に行うものである。また、スキーム別評価とは、外務省が持つ援助形態（スキーム）のうちのひとつの形態を対象に評価するものである。なお、経済協力局評価室は2003年3月31日付で廃止され、その機能は目的別に大臣官房考査政策評価官室と経済協力局調査計画課とに分割され、経済協力局評価室長は、大臣官房考査・政策評価官になった。

位レベルの評価における十分な準備・予算の手当が必要である。

このように、ODA 評価に関する議論においては、政策レベル、プログラム・レベルでの評価の可能性をめぐる技術的に難しい課題に直面したのである。そしてその答えの中では、プログラム・レベルの評価とは何かという問題に対する答えの輪郭は見えたと、「プログラムとは何か」ということまでは言及できなかった。

2.2 国内行政における政策評価の動き

「行政機関が行う政策の評価に関する法律」が2001年6月22日に参議院を通過、成立したことは、わが国政府が先進国並みの制度として評価制度を持つに至った象徴的事件である。この法律によって、それまで公共事業をはじめとした個別事業評価(プロジェクト評価)だけでなく、政策レベルでの評価についても本格的な取り組みが進んだからである。

ところが実際に実施してみると、政策評価には重要な思考方法が3つ必要であることが次第に明らかになってきた。すなわち、目標 手段という因果関係的な視点から政策を「政策 施策 事業・事務事業」というように体系化した政策システム概念(これをロジック・モデルとかヒエラルキー・モデルと呼ぶこともある)、「インプット アウトプット アウトカム インパクト」という流れのプロセス概念(この考え方をリニア・モデルと呼ぶこともある)そしてこの政策システムやプロセスをうまく整理し、適切に動かす概念装置、すなわちプログラムである。後述するように、プログラムは活動そのものというよりは、活動を導くアイデア、ソフトに近い。しかし、現場ではこうした点に思いが至らないためにいくつか問題が発生した。たとえばプログラムが本来の機能を果たしておらず、違った意味で使われている(地方自治体の男女共同参画プランや健康21プラン)あるいはそもそも存在しない(多くの市町村の場合)という問題である。

その結果、地方自治体の場合、政策評価とは言いながら実際には事業評価、業務の見直し、行政評価という名のコスト削減策が主流である。施策評価と呼んでも、それは事業より規模と予算

が大きい活動の評価であったり、あるいは総合計画や基本計画でとりあえず「施策」と呼ぶものの評価になっている。しかも、これら施策は事業と異質のものではなく、明確に区別できる特徴はなく、そもそも施策が何であるのか説明していないことが多い。

同じことは国の府省についてもいえる。実は政策評価を導入するために開催された旧・総務庁行政監察局の「政策評価の手法等に関する研究会」でも、第14回の会合(2000年6月19日)の中間まとめ素案検討段階までは、政策評価の代表的な方法として「政策体系評価」、「施策実績評価」(強調は筆者)、「事業評価」という形で検討されていた。たとえば政策体系評価に関して集中的に討議した第11回会合(2000年4月25日)では、政策体系評価は施策と関連付け、次のような視点で情報を産出するものであると考えられていた。

- ・施策等が所期の目的、目標を達成し、効果を上げているかを明らかにする。
- ・施策等の目的が実現される因果関係や波及効果のプロセス、施策等の寄与度、外部要因の影響等を分析する。
- ・施策等の問題点を発見する。また、その原因を解明する。
- ・施策等の必要性を検討する。
- ・施策等による効果や負担を把握し、効率性を検討する。
- ・施策等の改善や見直しについて提案する。

その際、政策体系評価はアメリカ合衆国会計検査院(GAO)のプログラム評価との類似性が指摘されていた。また、施策実績評価とは、これもアメリカで1993年に制定された「Government Performance and Results Act」の議論と重ね合わせて論じられており、明らかに業績測定(performance measurement)を意識していた。

こうした考えの前提となるべき概念整理が、政策と施策、事務事業の定義として「中間まとめ素案」の討議(2000年6月19日の第14回会合)で示されていた。すなわち、「『政策』は、特定の行政課題に対応するための行政活動の基本的な方針を示すものであり、この方針の実現という共通の目的を持った行政活動の大きなまとまりと捉えることもできる。」「『施策』は、上記の『基本的な方針』を実現する為の具体的な方策であり、この方策の実現という共通の目的を持った

行政活動のまとまりと捉えることもできる。」
『事務事業』は、上記の『方策』を実現するための個々の行政手段としての事務または事業である。このような事務や事業は、行政活動の基本的な単位となるものである。」

ただし、このような区別はひとつの理念型であり、3つに分かれない場合もあり、また社会経済情勢の変化に対応して変化するので、その区分についてもあまり固定的に捉えるべきではない、ということもわざわざ注記している。逆に言えば、中央府省では政策と施策、事業の関係が必ずしも明確でないことを告白しているのである。

そして、6月27日に開催された第15回会合では、政策体系評価は「総合評価」に、また施策実績評価は「実績評価」に名前を変えた。それぞれの名称については、評価の対象と評価の手法の名称とは別の話である、施策や事務事業についての説明は理解が難しいなどさまざまな議論があり、その結果、最終的に名称については村松岐夫座長に一任し、変更が認められた経緯がある。つまり、施策やプログラムについての定義は、わが国政府の公式的な統一見解としては存在しないのである。それでは理論研究においてプログラムはどのように定義されるのであろうか。

3. プログラム概念

3.1 一般的な説明

プログラムそれ自体の定義として有名なのはマーチとサイモンの『オーガニゼーションズ』である。彼らは一般論として、プログラムの概念を組織活動の内容という視点で説明する³。「ある状況の下では探索と選択の過程が非常に短縮化する(中略)。その極限では、環境からの刺激が直ちに高度に複雑で体系化された反応の集合を、その組織から喚起する。このような反応の集合を、われわれは実行(performance)プログラムないしは単にプログラムという。たとえば消防署

では、警報が鳴るとこのようなプログラムが始動する。ソーシャルワークのデスクに生活保護者が現れると、(中略)同じようにプログラムが始動する。」

マーチとサイモンは特定の組織がどのようなプログラムを用いているか判断する方法のひとつとして、標準作業手続が記述されている文書を検討する方法があるという。また、個々の活動の反復性が大であれば大であるほど、プログラム化はより大になるという。それを踏まえて彼らはプログラムの内容として、次の3つの命題が重要であると指摘する。

- a. 機械や作業の一引用者速度を決めるルールがプログラムの中に組み込まれている程度。
- b. 仕事の活動がプログラムの中で詳述されている程度。
- c. 製品の設計明細が、プログラムの中で詳述されている程度。

つまり、プログラムとは組織内で仕事をする成員をコントロールするシステムの一部であり、調整システムの重要な部分なのである。

さて、評価の領域でもプログラムとプロジェクトの説明は古くから存在した。すなわち、1970年に刊行された‘Federal Evaluation Policy’という報告書である⁴。ここでは次のようにプログラムについて説明する。アメリカの連邦政府が「プログラム」によって資金を提供し、現場(州、市、学区、地方住宅局、病院など)の活動の方向性を示す。現場はそのプログラムの指示に従って、予定された受益者にサービスを提供する。このサービス提供活動の単位が「プロジェクト」である。似たようなプログラムとプロジェクトの関係の定義は、1999年のOECDの報告書にみられる⁵。ここでは「プログラムとは一定のタイムスケジュールと予算に従って、政策を実施するために行われる一群の組織化された活動、たとえば期待された政治目標を達成するための条件を作る活動」であるという。

³ J. G. マーチ・H. A. サイモン(土屋守章・訳)『オーガニゼーションズ』、ダイヤモンド社、1977年、215～220ページ。

⁴ Joseph S. Wholey, John W. Scanlon, Hugh G. Duffy, James S. Fukumoto, Leona M. Vogt, *Federal Evaluation Policy: Analyzing the Effects of Public Programs*, The Urban Institute, 1970, p.24. ちなみにこの本は1968年に当時の合衆国連邦政府の住宅・都市開発省の委託を受け、当時の連邦政府において評価が可能であるかどうか、あるいはその能力はあるのかどうかという問題に関してシンクタンクである「アーバン・インスティテュート」が行った研究の報告書である。

⁵ OECD, PUMA (<http://www.oecd.org/puma/>), *Improving Evaluation Practices: Best Guidelines for Evaluation and Background Paper*, 1999, p.12.

3.2 テキストやマニュアルの「プログラム」

実務と理論が近接している ODA の分野では、国籍や社会背景の異なる人びとの協働が不可欠であるため、理論において言葉の定義はかなり熱心で、プログラムに関してもたびたび定義が試みられている。たとえば、「特定の(ただし通常は類似の、ないし関連する)目的を達成するために意図された複数のプロジェクトないし一連のサービス」というノルウェー外務省の包括的な定義がある⁶。また、より詳しい世界銀行のプログラム評価に関する研究報告書では、プログラムを「あるセクター中のプロジェクト群や諸活動のシリーズである。それらはすべて同じ目標を達成しようとしたり、同じサービスを提供しようとする。たとえば、特定の国や地域内での栄養補給プログラムというものがあるとすれば、それは国家予算から資金を得た活動だけでなく、多くのプロジェクトからなっている」と説明する⁷。ちなみに、ここで言うプロジェクトとは投資、政策、制度上の活動などの個別具体的なパッケージで、特定の開発目標を達成することを企図するものである。そしてこのプロジェクトが生産するものは直接生産物であるアウトプットであり、政策やプログラムの評価が対象にするアウトカム(成果)ではないという点に明確な特徴がある。

別のテキストでは、プロジェクトは特定された時間枠、人材、財政資源や物的資源を伴う目標達成の為に計画された介入行為であり、プログラムはプロジェクトと比べるとかなり曖昧で、包括的で、期間はプロジェクトよりも長期にわたり、多様な介入行為だと説明される⁸。そして両者の違いをより明確化するため、プログラムは‘strategic planning’によって形成される、という説明を付け加えている。その形成方法の手順は、問題分析‘problem analysis’、利害関係者分析‘stakeholder analysis’、目標分析‘analysis of objectives’、影響を及ぼす外部要因分析‘analysis of external influencing factors’、組織の

構造、能力、他の組織に代替できる可能性の分析‘analysis of responsible organizations’、の順に行われる。つまり、こうして分析した結果を言葉にして、仕事(プロジェクト)の進め方として文書化しているものがプログラムだということになる。

3.3 プログラム概念を活用しないわが国の事情

さて、わが国では実務の上で施策の概念は存在するかもしれないが、それを明確に定義したり、抽出することが困難である。理由はいくつか考えられる。

第1に、道路課や住宅局、教育委員会などのように組織がプロジェクト単位で縦割り化されているということである。プロジェクトの具体的な仕事名が組織名称になっているのである。したがって、道路の渋滞問題であれば道路課、小学校の治安・安全問題であればまず教育委員会というように、非常に安直な思いこみの中で「プログラム」に関する了解が成立し、それについては疑問を持たずにきた。つまりプログラム設計の適切さの是非を論じないだけでなく、そもそもプログラムを意識する必要がなかったのである。したがって、プログラムの再考を促す問題提起、たとえば「交通渋滞問題は道路課というよりは、住宅、公共交通などの総合的な都市開発担当部局の問題ではないか」とか、「クルマ社会を容認する個々人のライフスタイルの問題かもしれない」という提案は無視されるのである。

第2に、政策によってはプログラムの合意があり、しかも現場でのオペレーションのテクニクはそれなりに充実し、結果としてプログラムとして改めて意識することは不要だったという説明も可能かもしれない。戦災復興期、高度経済成長期までの哲学は「開発主義」であり、これに基づき現場ではプロジェクト事業として道路やダム、農地改良、学校や病院などのインフラ整備は行なわれる。これが開発のためのインフラ整備「プログラム」という暗黙の前提となり、

⁶ 国際協力事業団企画部評価監理課『ノルウェー外務省作成・開発援助の評価 評価のためのハンドブック』、1996年3月、118ページ。

⁷ Joseph Valadez and Michael Bamberger eds., *Monitoring and Evaluating Social Programs in Developing Countries*, The World Bank, 1994, p.471.

⁸ Reider Dale, *Evaluation Frameworks for Development Programmes and Projects*, Sage(New Delhi), 1998, pp.20-23.

ある時期までは問題にならなかった。プログラムが問題になるのは、この開発主義が社会情勢や時代背景とミスマッチを起こした時期（1990年代）である。

なお、とくに公共事業に関していえば、仕事の分担関係を定める組織体制が「政策と事業」というように分業体制としてセットされていた。たとえば省 事業団・ 公団、本省 地方支分部局、中央省庁 地方自治体である。この分担関係の中に、表面化せずプログラムが潜んでいる可能性はある。たとえば、かつて中央省庁から地方自治体に向けて出され、現在は廃止された「通達」はプログラムと呼べるものかも知れない。

第3の理由として、これまでの日本ではインフラ整備の公共事業を、景気対策との組み合わせで進めてきたという経緯を指摘したい。そして、これについては誰もが議論の余地が無いと思ひこみ、あるいは思考を停止しているので、細かな説明は不要であった。インフラ整備は景気対策という公然の含意を持つプロジェクトとして展開され、いずれが本来の目的なのかということについては疑問が出されなかったので、インフラ整備が所期の目的を達成したかどうかは問題にならない。したがって仮にプロジェクト目的が時代遅れで陳腐化し、全く無意味だということが公然の事実になったとしても、政策目的とその手段の関係をプログラム問題として改めて議論する必要は認められなかったのである。

このようにして、わが国ではプログラム概念にあたる思考方式が定着しないまま中央省庁改革と地方分権改革が進み、その一方で政策評価が登場し、施策という古くから存在した言葉をプログラムに当てはめてしまった。しかし、いざ評価対象として考えるとき、実は実際に評価できる対象としての施策・プログラムが政策の中には存在しないのではないかと、という困った事態になっている。

4. プログラム評価

4.1 プログラム評価の展開

プログラム評価という概念はわが国では一般化していないが、アメリカそしてOECD加盟諸国では‘social program’の評価という形で普及したこともあって、政策学・政策科学や監査・監察の分野ではなく、‘social service’や‘human service’に関わる学問分野、たとえば教育、福祉、医療、労働の分野で一般化し、また実務での活用が見られる。これらは‘social program’あるいは‘social service program’の見直しということで、多くのテキストも出版されている（たとえば出版社‘Sage’は、古くから多くの‘program evaluation’、‘evaluation research’のテキストを刊行している）もともと、わが国では‘social service’や‘human service’の分野でもプログラム評価とは何かという議論はあまり見られなかったし、不十分であり、なおかつその必要性も認識されてこなかった⁹。

しかしながら、プログラムの評価は存在している。既述した外務省のプログラム・レベル評価、国土交通省のプログラム評価、地方自治体の政策評価システムの施策評価があり、その他財団法人・日本建築防災協会「耐震診断プログラム評価」、全国8国立大学工学部の工学教育「プログラム評価」分科会、大阪産業再生プログラム評価委員会（大阪府商工労働部商工労働総務課）、宇宙環境利用研究委員会のプログラム評価（宇宙開発事業団）、「語学指導等を行う外国青年招致事業」のプログラム評価懇談会、日本政策投資銀行の「プログラム評価」など、数は増えている。ただ、こうしたプログラム評価の考え方を導く理論は詳しく展開されてきたわけではない。例外的に田辺国昭教授の論考がある¹⁰。

田辺教授は政策評価には行政活動評価、プロジェクト評価、そしてプログラム評価の3種類があるという。第1の類型である行政活動評価は、「行政組織内部で営まれている事務の効率化とその事務に従事する行政官のインセンティブの提供が主たる目的」である。その「評価手法としてはあらかじめ事務ごとに目標を数値化し、この数値をどの程度実際に達成したのかを見ることになる」。第2は「政策の枠組みを所与としてその下で営まれる個々のプロジェクトや規制

⁹ 例外的に社会福祉の分野では、早くから政策評価、プログラム評価の重要性について認識されてきた。たとえば、『社会福祉研究』第33号、特集テーマ「社会福祉実践の評価方法・意義・内容・課題」-、財団法人・鉄道弘済会、1983年10月がある。

¹⁰ 田辺国昭「政策評価の仕組み」、ジュリスト、No.1161、1999年8月1・15日号、148～153ページ。

を対象とする評価である。これらの評価においては、まず事業や規制を実施することによって、社会経済をより望ましい状況へと改善することができるか否かを明らかにすることが目的となる。」プロジェクト評価や規制のインパクト分析が主たる方法である。

第3の評価類型が本稿で議論するプログラム評価に相当するもので、これは「個別の事業ではなく、政策の枠組み自体を評価対象とする。」政策プログラム分析とも呼ばれ、「政策のさまざまな問題点を指摘し、その解決に向けての情報を提供することが主たる目的となる。」「政策の枠組みによって生じる問題や執行過程において生じる問題など、さまざまな分析手法を組み合わせることで、有用な情報を算出することが求められる。その結果、(中略) 定量化の占める比重は、必ずしも高くはなく、数量化しづらいさまざまな定性的な情報の占める比重が大きい。」「ここで産出される評価情報は、政策において何か着目しなければならぬ事態が生じていないか注意を喚起し、またそこで明らかになった問題を解決する為に必要な情報であり、政策決定者にとって次の行動を起こすためのよりきめの細かい情報を提供することになる。」

いささか引用が長すぎたが、この説明で見る限りわが国の中央府省で導入された政策評価の3方式のうち、「実績評価」は田辺教授の行政活動評価に近く(田辺教授は「政策評価の手法等に関する研究会」の第14回では施策実績評価は「行政活動評価」もしくは「業績評価」という呼称がふさわしいのではないかと述べている)「総合評価」はプログラム評価とほぼ重なると思われる。

なお、前述した報告書『Federal Evaluation Policy』においては、プロジェクトの評価とプログラムの評価とは明確に区別されている¹¹。すなわち、プログラムの評価としては‘program impact evaluation’と‘program strategy evaluation’がある。前者は、ひとつのプログラムがその全体的な目標達成するときの有効性を、あるいは複数のプログラムがそれらの共通の目標を達成する際の実効性を評価するものである。インパクトを測定するためにアウトプット変数や比較対照法が使われる。現在わが国で導入されている政策評価が「施策の成果」を検証するという場合、この

考えに近い。もう一方の‘program strategy evaluation’は、ひとつのプログラムに使用するさまざまなテクニック、手法の有効性を評価し、比較する。主に、当該プログラムのプロジェクトで使っているさまざまな戦略や手法の有効性を、それぞれ比べてプログラム・マネージャーに情報提供するために行われる。これはアメリカ合衆国で‘social program evaluation’に使用される。

これに対してプロジェクト評価としてはふたつ考えられている。ひとつは個々のプロジェクトの目標達成度の評価でいわゆる‘project evaluation’であり、事業評価(公共事業評価)である。他は‘project rating’であり、異なる地域で実施される複数のプロジェクトがそれぞれのプログラム目標を達成するときの有効性を比べようというものである。一個のプログラムの中で地域ごとに複数のプロジェクトが運用されているが、それらの成否を評価し、それを比較した情報をプログラム・マネージャーに提供しようという方法である。この後者の‘project rating’は、実は本稿のプログラム評価のイメージに近い。

4.2 国土交通省の「プログラム評価」

ところで、中央省庁のうちで国土交通省は明確に「プログラム評価」という概念を打ち出している。2001年7月27日、総務省の政策評価・独立行政法人評価委員会における国土交通省の政策評価担当者の説明によると、国土交通省は3つの政策評価の方式を採用している。新規施策について必要性、有効性、効率性をチェックする事前評価の「政策アセスメント」、目標を具体的な指標で示しその達成度を測る業績測定タイプの「政策チェックアップ」(他の府省では実績評価とも呼ぶ)そして総合的な視点で、掘り下げた分析・評価を実施する「政策レビュー」である。政策チェックアップは定期検診・定期健康診断のような機能を担い、その検査に引っかかった問題政策についてより深く掘り下げた調査・見直し改善をするために政策レビューが行われるのである。政策レビューとして実施されている例はダム事業、都市圏の交通渋滞対策、国際ハブ港湾のあり方、土地の有効利用、海洋汚染に関す

¹¹ Federal Evaluation Policy, op.cit., p.25.

る取り組みなどである。そして、この政策レビューを国土交通省はプログラム評価と呼び、特定テーマに関連する施策群(プログラム)を対象として、プログラムの実施と結果の因果関係を詳しく分析し、所期の効果を上げているかどうかについて検証し、課題と改善方法を発見する手段として使っているのである。このプログラム評価の考え方は、中央府省全体の「政策評価に関する標準的ガイドライン」(平成13年1月15日、政策評価各府省連絡会議了承)における総合評価に近い方式である。

4.3 県レベルの「施策評価」

地方自治体では「政策 施策 事業」という体系を作り、政策評価システムの重要な構成要素のひとつとして施策評価を置く例が一般化している。ここでは地方自治体のすべてを取り上げる余裕はないので、筆者が関わった岩手県と秋田県の政策評価を例にとって、そうした自治体の政策評価システムの論理構造を明らかにしたい。

2001年度から実施されはじめた岩手県の「政策評価」(当時、総合政策室政策評価課が担当)は、岩手県総合計画を前提とする評価システムであった。ただ、総合計画は政策評価導入以前に完成しており、評価を前提とした政策体系(政策 施策 事業)とは仕組みが違っていたので、総合計画を政策体系として働くオペレーショナルな仕組みに再編成するために、かなりの時間と労力を費やした。

再編成の結果、この総合計画が目指す「5つの社会」、すなわち「自然と共生し、循環を基調とする社会」、「快適に安心して暮らせる社会」、「創

造性あふれ、活力みなぎる産業が展開する社会」、「ネットワークが広がり、交流・連携が活発に行われる社会」、「個性が生かされ、共に歩む社会」の5つの社会を実現する手段として17の「施策」を置き、総合計画を政策評価のために操作可能な概念体系に組み替えようとした。

具体的には、たとえば「自然と共生し、循環を基調とする社会」には「参加と共同により環境にやさしい地域社会の実現、人と自然が共なる環境の保全、循環型社会の創造、という3つの施策が置かれている。また、その施策にさらに「分野」という概念装置がつけられ、17施策は合計78の分野を展開することになる。すなわち、この施策においては3種類の分野、「地球環境問題への地域からの取り組み」分野(グリーン購入ネットワーク会員数が業績達成の指標)、「環境を守り育てる参加と協働の仕組みづくり」分野(ISO14001認証取得団体数が指標)、「環境について自ら考え、行動する人づくり」分野(環境アドバイザーの研修会への派遣、子供エコクラブ体験学習支援)という分野が置かれているのである。枠組みだけ説明してもこのように複雑であり、評価は非常に大きな労力が必要になってくる。ちなみにここで言う施策が政策にあたり、分野が施策、主要事業と細事業が事業に相当し、他の地方自治体と比べると混乱する用語法になっていた。

岩手県の政策評価は、作成しなければならない評価調書が施策評価調書(施策の進捗状況を総合的に分析)、分野評価調書(分野の総合的な達成状況を測定、県民意識の変化や社会経済情勢を分析)、主要指標評価調書(主要な指標の達成状況を測定し、主要事業の有効性などを分析)、主要事業評価調書(主要事業の達成状況を測定し、事業効果等を分析)と4種類もあったた

表1：岩手県政策評価システムの骨格(2001年当時)

5つの社会 17の施策 78の分野 678の主要な事業(2001年度)

この78分野には195の主要な指標が置かれていた。

表2：秋田県政策・事業評価システムの骨格(2000年当時)

21の政策 70の施策 186の数値化した施策目標 施策目標と関連した事業

この他大規模事業について、特別な評価手法が置かれている。

め、さらに混乱を招いた。

他方、秋田県では「政策・事業評価」を1999年度から運用しており（企画振興部総合政策課が担当）、2000年2月に策定した「あきた21総合計画」は政策評価導入後に策定されたことになる。したがって、岩手県とは対照的に、秋田県総合計画は政策評価と親和性があり、「政策 施策 事業」というシステムが一貫性を持つ分かりやすい仕組みになっている。

秋田県の政策・事業評価システムの要点は、そこで作成される評価カードを見ると明快である。すなわち、定性的な目標設定、関連する施策の目標達成度、重要度、緊急度、方向性などの総合評価を記述する政策評価カード、施策目標達成度を記述する施策評価カード、「あきた21総合政策」で設定した186の指標について書き込んだ施策目標別評価カードの3種類のカードを作成することになっていた。事業評価も別途行っており、ここではそれぞれ様式を定めて県関与の妥当性、対象・手段の妥当性、目標達成度、費用対効果分析、住民満足度、施策目標への貢献度を評価し、最後に総合判定（継続・改善・縮小・休廃止）を行うことになっていた。なお、2002年4月に「秋田県政策等の評価に関する条例」が制定された後、この仕組みは一部変更になっているが、政策と施策を評価の中心とする方針は変わっていない。

秋田県と岩手県の政策評価を比較すれば、結局は総合計画が政策評価を前提として作られているかどうかという問題になる。総合計画が意識的に政策体系との親和性をはかっているシステムも分かりやすく、評価を行うための制度構築作業はより簡単になる。親和性が無く、総合計画と政策評価システムが、全く別々に作られていれば、総合計画のシステムを政策評価に合わせて組み直す必要上、どうしても複雑になる¹²。

しかしいずれにしても、岩手県や秋田県だけでなく多くの地方自治体は、政策体系を前提としてはいるものの、基本的には実績評価・業績測定を使用している。そして、いずれもアカデミズムや諸外国の評価機関が使うプログラム概

念とは微妙に食い違った施策概念を使用している。つまり、評価対象として見るとわが国のプログラムはまだ実用が難しい。そこで政策と、その手段である施策や事業との因果関係の検証がないまま、仮に置いた施策目標なるものの達成度を測定する作業が続けられてきたのである。結果としては、有効なプログラム・プロジェクトと無効なものの選別ができず、単なる効率や業績の測定に終始し、せいぜいコスト削減の理由付けだけが可能な評価になってしまい、それに対して現場から徒労感と失望が高まったのである。

4.4 プログラム評価と業績評価、モニター

このような地方自治体の政策評価に対する徒労感と失望は、プログラム評価、モニター、業績測定の3つにおける方法と目的の違いを明確に区別しないまま導入した結果発生したと考えられる。

とくに、多くの地方自治体で導入した業績測定の方式と数字をモニターする方式は、それが良いか悪いかは別として、行政活動の目的に指標を付けその指標の達成度を測定すれば可能である。つまりその対象が施策・事業のいずれであっても可能であるため、政策と政策手段（施策・事業）との区別は不要である。「評価対象は何か」という議論で紛糾しないのである。しかも「施策評価」という名称の評価は数多く存在する。この結果として、多くの地方自治体ではプログラム評価と業績測定、モニタリングの区別は曖昧なまま放置されている。しかし、現実にはこれら3つについては明確な区別がある（表3を参照）。

とくに業績測定、モニタリングのふたつは、行政の内部管理情報を産出する上では意味があるが、住民に対するアカウントビリティ、政策の効果が住民にいかなるインパクトを持っているかという面での情報を産出する方法ではない。政策目的を達成する手段を改善するために必要な情報は得られない。測定やモニターの対象になる数値、指標は一種の「めやす」にすぎないことが多く、政策の成果・インパクト・政策目標とその手段の適切な関連性を証明するものではない。したがっ

¹² 総合計画や基本計画については、もっぱら「住民をいかに参画させるか」という視点が重視され、さまざまなテクニックが開発されてきた。しかし、その参加のあり方を検証すること、あるいは総合計画・基本計画の評価についてはあまり議論されていない。こうした問題点に関しては、山谷清志「自治体計画の実施・評価と住民」、西尾隆編著『住民・コミュニティとの協働』、自治体改革9、ぎょうせい、2004年を参照。

表3：業績測定とモニタリング、プログラム評価の比較

	業績測定	モニタリング	プログラム評価
分析の単位	プログラム・プロジェクト	プログラム・プロジェクト	プログラム
主たる目的	外部への報告	プログラム管理	プログラムと政策の改善
視点	財務管理	マネジメント	政策とプランニング
データの利 用方法	外部の利害関係者に対し、プログラムを運営した結果の業績に関する情報をフィードバックする	実施機関のマネージャーに対し、プログラム運営に関する情報をフィードバックする	プログラムの成果（アウトカム）とインパクトに関する情報を、政策作成者やプランナーにフィードバックする

出典：Peter M. Kettner, Robert M. Moroney, Lawrence L. Martin, *Designing and Managing Programs : An Effectiveness-Based Approach*, second edition, SAGE, 1999, を一部修正して引用。

て、政策作成者やプランナーに対する情報フィードバックとしては、プログラムの中身、内容を評価する方がより有益な情報を提示しやすいし、また、サービスの顧客に合わせてプログラムやプロジェクトのあり方を改善するにもプログラム評価が役に立つ。

ただ、プログラム評価を着実にこなすためには、「評価（evaluation）」が応用社会科学（applied social sciences）と呼ばれるように社会学、心理学、統計学など多くの学問分野を評価の実務、すなわち情報の収集、調査、分析、比較、そして評価結果の見せ方の実務に応用するノウハウが不可欠である。しかし、わが国では実務も研究も、シンクタンクですらそうした要請に応えるには十分でなく、まだまだ研鑽の必要がある¹³。そして、プログラムの概念が曖昧なまま評価システムが運用されてしまったのである。政策評価が重視する政策立案、政策決定の場へのフィードバックは期待できそうもなかったのである。

4.5 プログラム評価とプロジェクト評価の比較

ここまでのプログラムとプロジェクトの議論を整理するためには、実務で一般化しているプ

ロジェクト評価（事業評価）と、わが国ではまだ開発途上にあるプログラム評価（総合評価・国交省の政策レビュー）とを比較することが有益である。そして両者の明確な違いとして、次の4点が指摘できるであろう。

第1にプログラム評価の方が、時間的にも、空間的にも、また問題や課題にアプローチするコンセプトの上でも、プロジェクト評価よりトピックを広く、深く掘り下げて扱う。

第2に、受益者間や活動単位間で付けられる優先順位は、プログラム評価の方がプロジェクト評価よりも大きな規模で、広範囲に行われる。これについては、いくつか実例が出はじめている。たとえば2000年当時秋田県総合政策課が行った議論では、事業を特定の箇所（場所）で評価するのではなく事業総体として評価しようと試みた。すなわち、都市計画街路事業に関して特定の工事箇所ではなく、9カ所で行われている事業を総合的に見ようとしていた。この場合、個々の地域で行なわれている街路事業の進捗状況、費用対効果を見る場合はプロジェクト評価であるが、9カ所の工事全体として街路事業を見た場合、それはプロジェクト群の評価ではなくプログラムの評価と認識した方がよいかも知らない。

あるいは、高齢者障害者移動円滑化法（通称・

¹³ 学問を政策評価の実務に応用するときの課題については、山谷清志「政策評価とシンクタンク 不幸な出会い」、NIRA 政策研究、2001年、vol.14, no.4, 特集『シンクタンクの方向性と政策研究』を参照。

交通バリアフリー法)はさまざまな施設についてバリアフリー化を求めるものであるが、これもまた、広く各事業横断的な視点でその取り組みを点検するという意味でプログラム評価の範疇に入れられるであろう。

第3にプログラム評価では、調整(co-ordination)の視点が重要であり、この点がまさにマーチとサイモンのいうプログラムの機能であって、こうした機能をプログラムが十分に果たしたかどうかプログラム評価では重要になる。たとえば、中央府省から地方自治体に補助金で仕事が行われている場合、あるいは援助国から被援助国に資金協力で仕事が行われた場合、中央府省の政策と地方自治体独自の政策との調整、あるいは援助国の政策と被援助国が考えている政策の調整を考えると、プログラムの存在とそれを評価する制度は重要である。中央府省の下請け機関として自治体があるわけではないし、被援助国がすべて援助国の政策フレームの枠内でプロジェクトを展開しているわけではないからである¹⁴。

第4に、プログラム評価は特定の争点(たとえば土木や医療、あるいは金融などの技術上の問題)、プログラムの各パート(内部にあるプロジェクト群)には、あまり関心を持たない。公共事業評価の議論が、ともすれば土木工学や生態系など理科系の専門に集中するのは、それがプロジェクト評価だからである。インフラ整備を経済分析する話も同じである(したがって経済学部のカリキュラムにはプロジェクト評価論がある)。プログラム評価は特定のプロジェクトが所与なのではなく、異質なプロジェクトが複数存在する中で、それらの間での選択、あるいは優先順位を付けるときに使われ、またその選択と優先順位の適正さを検証するために使用される。

5. むすびにかえて

本稿の問題関心の根本は以下のようなものであった。すなわち、現場で行なわれている事業や業務が「良い仕事をした」と評価されるかも知れないが、それだけでいいのかどうかということである。たとえば、男女共同参画社会の構築のた

めに立派な女性センターを建設し、その施設で講演会を開催したところ非常に盛況であった場合がそれである。立派な施設ができたこと、講演会が盛況だったことと、男女共同参画社会の実現とは必ずしも直接の因果関係はない。また過疎対策としてホテル建設、高速道路の整備をしても、過疎問題の本質がそこに住む高齢者の日常生活問題(病院への往復、買い物、屋根の雪下ろし)であるとき、この過疎対策はプログラム・ミスであろう。

逆に、農水省のある補助金のように、市町村に平等にばらまくのではなく、地域社会における女性の社会進出(農協の役員、各種委員会の委員、家庭内で家事を平等に負担する「憲章」づくりなど)が進んでいると見なされる市町村に優先的に補助金を配分する仕組みを組んだときには、男女共同参画社会の実現は近づくであろう。本音では女性の社会進出に消極的であっても、補助金が欲しい場合にはそうせざるを得ないからである。これはプログラムがうまく機能した例になるかもしれない。

そうした意味で、プログラムとは制度や仕組み、プロジェクトを動かすマニュアル、アイデアである。児童虐待や薬物依存、道路渋滞などの課題をどのような方法で解決するのか、そのノウハウであるかもしれない。そしてプログラム評価とは、そうしたアイデアやノウハウを対象とする評価であり、官庁のジャーゴンでいえば政策をうまく「回す」方法の評価なのである。そしてこれらアイデアやノウハウが評価という検証に耐えるように作られていなければ、プログラム評価はそもそも不可能である。わが国の政策評価で決定的に不足しているのは、この発想なのである。

¹⁴ この点に関しては、三好皓一「プログラム・セオリー・マトリックスの活用について - 評価における分析の有用性を高めるために -」、日本評価学会『日本評価研究』、Vol.2.No.1, June, 2002、を参照。

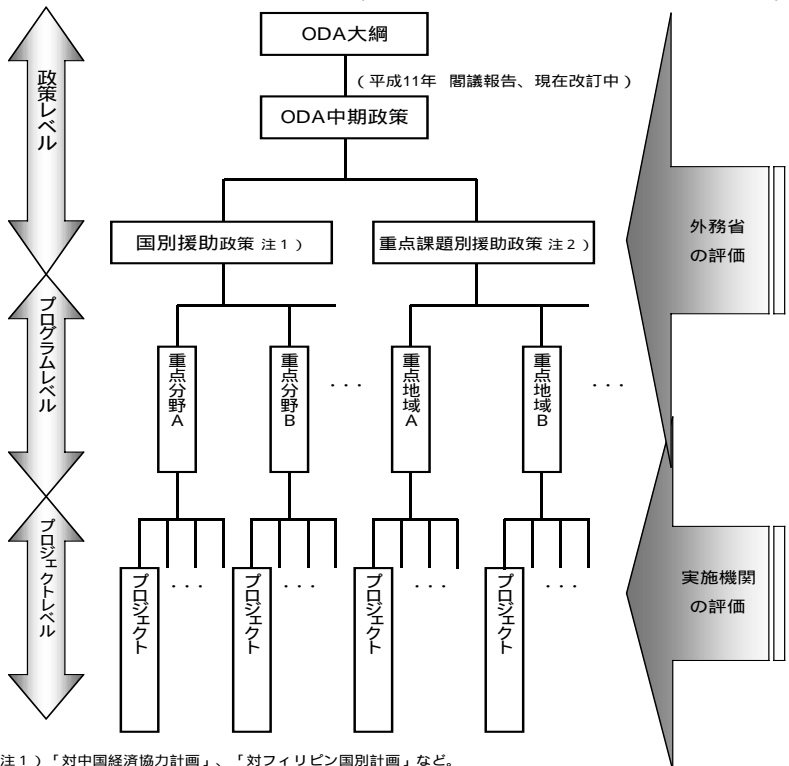
ODA 評価略年表

- 1981年 外務省「経済協力評価委員会」の設置。海外経済協力基金、評価の専門部局を設置。国際協力事業団、評価検討委員会設置。
- 1982年 外務省経済協力評価委員会、『経済協力評価報告書』公刊。
- 1983年 国際協力事業団評価検討委員会、報告書「形態別終了評価のあり方」を作成。
- 1984年 外務省、経済協力局に「調査計画課」設置（総務班、計画班、情報管理班、評価班）。
- 1986年 外務省経済協力局長、私的諮問機関「援助評価検討部会」設置。
- 1988年 国際協力事業団、企画部に「評価室」設置。
- 1990年 国際協力事業団「評価室」、昇格・独立して「評価監理課」。
- 1997年 「21世紀に向けてのODA改革懇談会」外務大臣の懇談会として設置（4月）。
- 1998年 「21世紀に向けてのODA改革懇談会」、最終報告書、提書（1月）。これを受け、「援助評価検討部会」および「評価研究作業委員会」ODA評価体制の現状、問題点、課題、改善案等について検討（1998年11月から1999年11月まで）。
- 2000年 「援助評価検討部会」、『ODA評価体制』の改善に関する報告書」を取りまとめ、河野外務大臣に提出（3月）。「ODA評価研究会」設置（7月～12月）。
- 2001年 「ODA評価研究会」報告書提出（2月）。
- 2003年 外務省経済協力局評価室、廃止（3月）。

ODAの政策体系および評価対象

【資料】

（平成4年6月閣議決定、15年8月改訂版を閣議決定）



注1) 「対中国経済協力計画」、「対フィリピン国別計画」など。

注2) 「京都イニシアティブ」、「途上国の女性支援」など。