

第2章 政策評価の進め方

国土交通省では、以上のような基本理念に基づいて政策評価体系を組み立てた。ここでは、個々の評価方式の解説に入るまえに、それらが全体として何を目指し、どのように関係づけられるものであるかを説明する。

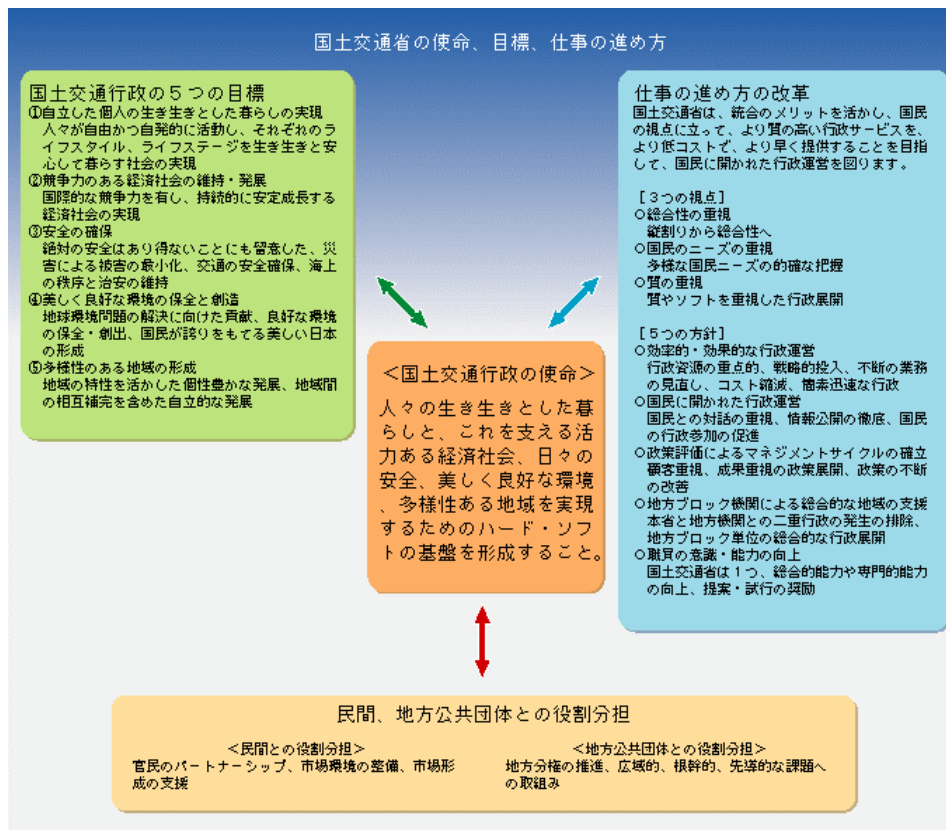
1 目標によるマネジメント

上記のような理念から、省や各部局の政策目標を明確にすることがまず必要である。全体の目標なしに各課が施策の見直しを行うだけでは、今ある施策の体系が前提となってしまうがちであり、社会の変化等に対応した機動的・戦略的な施策展開を講じることが難しい。組織全体の目標設定をまず行い、その目標達成のために何をすればよいかをそれぞれの部署が自律的に考えることによって、組織全体が同じ目標に向かった、一体となった業務運営が可能となる。こうした「目標によるマネジメント」(MBO: Management by Objectives)の考え方を行政運営の基本にすることが、上記理念から導かれる政策評価導入の大きな目的となる。これによって、部局を超えた施策の連携や新たな施策展開も促進される。

その際、今ある施策を前提として、それを束ねることによって目標を帰納的に定義するのではなく、国民にとって何がどう改善されるかというアウトカムを出発点として、そこから目標を演繹的に抽出することが重要になる(コラム2)。

先に述べた国土交通行政のビジョンの検討を踏まえ、国土交通省発足直後に取りまとめられた「国土交通省の使命、目標、仕事の進め方」(図2-1)は、省の使命(ミッション)や目標を示すものである。

(図2-1) 国土交通省の使命、目標、仕事の進め方



コラム2：アウトカム目標について

行政の目標にはさまざまなものがありますが、いずれも組織の使命（ミッション）を達成することが基本となります。通常はミッションでは漠然としすぎるので、政策分野ごとに、政策が目指す望ましい社会的状況や状態を記述します。これがアウトカム目標と呼ばれるものです。目標自体は定性的なもので構いませんが、目標によるマネジメントの考え方に立って、各政策企画立案部局が、目標達成に効果的な施策等を企画立案できる程度に内容が明確である必要があります。例えば、「地域活性化」という抽象的な言葉だけでは、十分明確な目標とは言えません。

以下は、成果重視の行政運営を積極的に推進していることで知られる米国アイオワ州政府が、職員向けに作成した「アウトカム目標“発見”フローチャート」です。

（施策名）の目的は、（施策の対象者）に対して（施策の内容）というサービスを提供し、彼らが（施策の目標とする状態）ことである。

（例）「就業促進施策」の目的は、福祉給付金対象者に対して、職業訓練の機会と職業斡旋というサービスを提供し、彼らが公的助成の対象から離れて経済的に自立することを支援することである。

（参考文献）State of Iowa (1998) Budgeting for Results Handbook.

また、欧米では、目標設定にあたって、以下の「SMART 原則」を満たすことが望ましいとされています。

Specific（具体的）

Measurable（測定可能）

Ambitious（意欲的）

Realistic/Relevant（現実的・上位目標との関連性）

Timed（時宜を得た）

2 業績測定

目標によるマネジメントでは、現場の裁量を認めるが、それだけでは組織全体のマネジメントがうまく進んでいるかどうかを確認することができない。そのため、戦略に沿った目標の達成度合いを測定することが必要になる。これによってはじめて、組織が目標実現に向けてうまく機能し、きちんと業績を上げているかどうかを、全体として俯瞰することができる。これを行うしくみが業績測定（Performance Measurement）である（コラム3）。業績測定は、目標によるマネジメントのための基本的なツールとして、民間企業や欧米の行政機関でひろく採用されているものである（コラム4）。

政策目標は、定性的な記述であるため、その達成度合いを把握するためには、それぞれの政策目標を代表する定量的な指標を設定し、その指標の変化を測定するという工夫が必要になる。このために設定する指標を業績指標という。このように、業績測定は、測定の難しい目標の達成度を、関連する指標を設定・測定することによって簡便かつ客観的に把握しようとするものである。したがって、業績測定を実施する際には、指標の選定が極めて重要になってくる（コラム5）。

業績測定は、具体的には、それぞれの指標に設定された目標値と、定期的に測定する実績値と対比することにより行うが、目標の達成度合いの「善し悪し」を判定することが目的ではない。

業績測定はむしろ、それを契機として、目標の達成手段としての関連施策の有効性を比較検討したり、目標が十分達成されていない場合に、その原因を外部要因の影響も含めて明らかにすることによって、成果についての行政の説明責任を果たすためのものである。また、成果重視の観点から、目標達成に向けた現場の努力をかき立てる動機付け（パフォーマンス・ドライバー）となるべきものである。

コラム 3：政策評価の歴史と業績測定

政策評価は米国において、主に医療、教育、国際開発などの分野で発達してきました。その初期の試みは、社会実験などの分析手法を活用して、実施したプログラムの効果を客観的に把握しようとするものでした。このような取組を通じて評価のための方法論が確立され、やがてその他の政策分野にも応用されていきました。これが**プログラム評価**の発達につながります。

ところで、政策評価発展の歴史を見ると、定量的な分析手法を用いて、できるだけ厳密な評価を行えば、唯一絶対の「正しい解」に至ることができるとする考え方が主流でした。しかし、現実の複雑な政策課題に対しては、定量的な分析手法は万能ではなく、定量的分析をあてはめることのできる事象はどうしても限られてしまいます。これに対して近年では、ことさらに定量的な分析手法を適用することにこだわるよりも、評価が全体として意思決定者や利害関係者の判断に役立つ情報を提供するものとなっているかという視点が重要視されるようになってきました。

業績測定は、意思決定者にとって有益な情報をタイムリーに提供することを目的とするものであり、このような流れに沿うものといえます。また、ベンチマーキングなどの企業経営の手法からも大きな影響を受けています。

業績測定は、米国の地方自治体などで 1960 年代から用いられてきたとされていますが、最近の急速な普及を先導してきたのは、米国ワシントン DC の政策シンクタンクであるアーバン・インスティテュートのホーリィ（Joseph S. Wholey）とハトリ（Harry P. Hatry）というふたりの研究者でした。ホーリィは、連邦各省に勤務した経験から、実験計画法などを用いた大規模な評価プロジェクトの結果が、政策立案者や現場の実施者が必要とするタイミングで提供されていないことに懸念を感じ、そのようなフォーマルな評価手法から離れて、業績測定のしくみを洗練させてきました。一方、ハトリは、行政活動を効率化しようとする自治体職員の取組を支援するため、行政マネジメントのツールとして業績測定の導入に取り組んできました。この両者が主導する形で、1980 年代に研究と実際への適用が進められました。

業績測定が一躍脚光を浴び、政策評価の代表的方式として認められるに至ったのは、1992 年に出版されたオズボーンとゲーブラーの著書“Reinventing Government”によるところが大きいといえます。同書の影響を受けて、1994 年に導入された米連邦 GPRA 法では、連邦政府の全機関に業績測定の実施が義務づけられることになりました（コラム 4 参照）。

（参考文献）佐々木亮・西川シーク美実(2001)「パフォーマンス・メジャーメント～その光と影～」日本評価研究 1(2)。

William R. Shadish, Thomas D. Cook and Laura C. Leviton (1991) *Foundations of Program Evaluation: Theories and Practice*, Sage Publication.

コラム 4：業績測定の実例～米国 GPRA

米国連邦政府では、「政府業績評価法」(GPRA: Government Performance & Results Act of 1993)に基づき、各省庁ごとに目標を設定して、定期的に達成度合いを測定する業績達成度評価が実施されています。

各省庁は、向こう3年間以上の省全体の戦略を定めた「戦略計画」(Strategic Plan)を作成し、各省庁が取り組むべき使命(Mission)を明示するとともに、達成すべき戦略目的(Strategic Goals)を策定します。さらに、この戦略計画を基に、年度毎の「実施計画」(Performance Plan)を作成して、その中で戦略目標の達成に向けた業績目標(Performance Goals)と業績指標(Performance Measures)とを設定します。各年度が終了すると、業績指標を用いて実績が測定され、年度内の業績目標を達成したか、戦略目的の達成に向けての推移が良い傾向にあるかなどが評価され、仕事の改善に活用されています。また、測定した結果はすべて、インターネットなどを通じて国民に公表されています。

例えば連邦交通省(DOT)では、「重要な国益を高め、かつ国民の生活の質を高めるような、安全な交通システムを確立する」という使命の下に、安全性、移動しやすさ、経済成長、人間と自然環境、安全保障という5つの戦略目標を掲げています。また、戦略目標2の「移動しやすさ」に該当する業績目標の1つ「交通へのアクセス性」では、「障害者に配慮されているバス車両の割合」や「障害者に配慮されている主要な鉄道駅の割合」を業績指標として設定するなど、省全体で61の業績目標、92の業績指標が設定されています。

(参考) DOT Strategic Plan <http://stratplan.dot.gov/>
DOT Performance Plan/Report <http://www.dot.gov/performance/>

コラム 5：望ましい業績指標の定め方

業績指標とは、業績測定をする上で、政策目標ごとに、業績の達成度を継続的に測定できる指標です。「道路の整備延長」などはアウトプット指標、「都市部のラッシュ時における平均走行速度」などはアウトカム指標と呼ばれます。欧米では、以下のような基準を満たすものが望ましい指標とされています。

米国開発庁(USAID)のチェックリスト

- Direct (直接的: アウトカムがダイレクトに反映されているか)
- Agreed (合意: 受益者、関係者が合意した指標か)
- Practical (実践的: 容易、低コストでデータが収集できるか)
- Reliable (安定的: 継続してデータが入手可能か)
- Objective (客観的: 測定者の裁量が入る余地がないか)
- Useful (実用性: 施策改善や意思決定する際に有用か)

米国行政学会(ASPA)による業績指標の選定基準

- 成果指向 (アウトカムに焦点をおく)
- 妥当性 (論理的にアウトカムから導き出せる)
- 変動の反映度 (業績の推移をうまく反映できる)
- 信頼性 (期間を通じて、同じ基準で情報を収集できる)
- 比較可能性 (異なる時点の業績を比較できる)
- 明確性 (一般国民が理解できる)

(参考文献) 龍慶昭・佐々木亮 (2000)「政策評価の理論と技法」多賀出版
米国行政学会(原著) (2001)「行政評価の世界標準モデル」東京法令出版

3 事前評価

業績測定は、目標によるマネジメントの観点から、導入されている施策全般が目標にそってうまく進んでいるかどうかをチェックするものであるが、これに対して、個別の施策を新規に導入する際にその必要性等を目標に照らしてチェックするのが事前評価である。事前評価は、政策分析（Policy Analysis）、政策アセスメント（Policy Assessment）などと呼ばれることもある（コラム6）。

新たに施策を導入しようとする以上、その必要性、効率性、有効性等を厳しく分析することは本来必要なことである。それを事前評価という明示的なプロセスとして行うことの意義は二つある。一つは、そのような分析の結果をオープンにすることで、政策の企画立案過程を透明にすることである。特に、目標によるマネジメントの観点からは、導入しようとする施策が省の目標や戦略にどう関連するものであるかを必ず明らかにすることが重要である。また、もう一つの意義は、施策の導入時にその意図や期待される効果を明らかにしておくことによって、事後に効果の検証を行う際の基礎情報とすることにある。これによって、事後に施策の効果を検証する段になって、その施策の当初の目的を「推測」とするような客観性に欠ける検証を防ぐことができる。

コラム6：事前評価の事例～英国規制改革

英国では、コスト負担増などの大きな影響を企業や慈善団体に与えるような、重要な法・規制案が議会に提出される際には、その担当省庁が事前に、「規制インパクト評価」（Regulatory Impact Assessment）を実施することになっています。ブレア政権の規制緩和政策の一環として実施されてきましたが、2001年規制改革法（Regulatory Reform Act of 2001）が成立して法制化されました。年間、約200本の評価が行われています。

規制インパクト評価は、その改革案が、省の目的などに照らしてなぜ必要なのか十分に説明した上で、企業などにどのような影響を及ぼすかを、可能な限り数値によって分析します。影響はプラスの影響（便益）とマイナスの影響（費用）とに分けて分析され、最後に、便益－費用によって、純便益が導き出されます。規制案だけではなく、その代替案の費用と便益も計算されます。費用の中では、規制を守るための費用（Compliance Cost）が重要視されています。例えば、担当部門を置いたり、書類の様式を変更したり、社内研修を行ったりする必要がある場合には、その費用がかかります。どんなに望ましい規制案であってもこの費用が膨大になると、企業などにかえって大きな負担となるため、事前に十分チェックされるのです。

規制インパクト評価の重要な点は、このような分析結果を用いて、利害関係者との協議（Consultation）が行われ、規制改革に向けた「合意形成」が図られることです。場合によっては、100以上の機関へのヒアリングが行われることもあります。評価結果をまとめたレポートには、誰とどのような協議を行ってどのような見解が得られ、それを踏まえてどのように規制案を修正したのかも、記述する必要があります。

交通・地方自治・地域省（DTLR）では、航空、建設、住宅、地域交通、鉄道、道路、海運などの様々な分野で、法規・規制の改革案に対して、この規制インパクト評価が実施され、その評価レポートはインターネット上でも公開されています。

（参考）交通・地方自治・地域省（DTLR） 規制インパクト評価ホームページ
<http://www.dtlr.gov.uk/pubs/ria/>

4 プログラム評価

業績測定は、目標の達成度合いを定期的にチェックして必要な場合に「警報」を鳴らす役割をもつ、例えば「人間ドック」のようなものといえる。これに対し、「精密検査」に相当するのが、プログラム評価（Program Evaluation）である。プログラム評価は、対象施策の効果の検証や改善方策の検討のために深く掘り下げた分析を行うもので、もともと GAO（米国連邦会計検査院）が連邦政府の施策（プログラム）に関して実施し、評価手法を開発してきたものである。また、欧米諸国では、GAO 以外でもひろく活用されている（コラム 7）。プログラム評価に係る手法は、何を解明しようとするかという評価の着目点によって、いくつかの種類に分けることができる（コラム 8）。

なお、ここでいうプログラムとは、ある政策目標を達成するための手段としての行政活動等の集合である。既存の施策、事業を個別に対象にすることよりも、共通の政策目標を持つ施策等を一括してプログラムとしてとらえ、評価対象とすることが望ましい。

プログラム評価は、時間やコストをかけて掘り下げた分析を行うものであり、あらゆるテーマについて毎年できるものではない。したがって、あらかじめテーマを決めて重点的、計画的に実施することが必要である。また、業績測定の結果と連動させ、目標を十分に達成できていない業績指標に係るテーマについてプログラム評価を実施することも考えられる。

コラム 7：プログラム評価の事例～カナダ

カナダ連邦政府では、財務委員会事務局（TBS）が 2000 年 3 月に発表した「カナダ国民に成果を」（Results for Canadians）において、連邦政府はカナダ国民のために存在するというマネジメントの基本的枠組みを打ち出し、市民指向、公共サービスの価値重視、成果重視の経営、責任ある公金支出、という 4 つの政策運営方針を示しました。さらに、2001 年 8 月には「成果志向の経営と責任の枠組み」（RMAF: Results-based Management and Accountability Frameworks）という、政策運営に実際に携わる現場マネージャーに対するガイドラインを作成しました。プログラム評価は其中でも、政策を深く掘り下げて分析し、どの仕事が成果の達成・向上に寄与しているのかを評価する方法として位置づけられており、RMAF 成功の鍵を握る重要なものとされています。

RMAF では、政策の実施・途中段階と事後段階の 2 段階でプログラム評価を実施することとしており、Relevance=政策は優先度を反映しているか、必要性・ニーズを満たすものか、Success=政策が意図した成果を実現しているか、反対に意図しない影響を与えていないか、そして最終的に実現すべき成果の達成に近づいているか、Cost-Effectiveness=成果の達成に向けて、代替案と比較して最も適切な手段を選択しているか、の 3 つの視点で評価が行われます。

例えば連邦交通省（Transport Canada）では、上記の方針のもと、5 年間で主要分野のプログラム評価を実施できるように計画を立てており、例えば「地方空港整備に関する補助プログラム」や「船員訓練・認証制度関連プログラム」など、年間平均 5～6 本のテーマを選定して評価を実施しています。

（参考）財務委員会事務局（TBS）評価ホームページ <http://www.tbs-sct.gc.ca/eval/>

コラム 8：プログラム評価について

米国 GAO によるプログラム評価の定義は、以下のとおりです。

プログラム評価とは、あるプログラムがよく機能しているかどうかを評価するための体系的な分析である。典型的には、プログラム目標がどのようにして達成されたか（あるいはされなかったか）、あるいは全体としてのプログラムの業績がどうであったかを詳しく分析する。業績測定結果を活用して、プログラムの効果や改善方策を検討することが基本となる。

ここでいうプログラムとは、「特定された目的や目標を持った行政活動、事業、機能、政策等の集合体」と説明されています。そして、GAO 等は、プログラム評価の典型的な手法として、以下の 4 手法をあげます。

プロセス評価

プログラムがどの程度当初の意図どおり実施されているかを評価する。典型的には、当初のプログラムで計画された質と量、あるいは客観的（第三者からみた）水準又は顧客の期待（成果）と実績とを比較する。（業績測定を実施する場合は、その結果をもってこれに代えることができる。）

アウトカム評価（セオリー評価）

インプット→アウトプット→アウトカムの因果連鎖を説明する。可能であれば、各インプットについてアウトカムとの相関性分析を行うこともある。事前評価時のセオリーを検証することもある。

インパクト評価

with / without 分析によりプログラムのネットの効果を導き出す。可能であれば、プログラムを構成する各施策等ごとに寄与度を分析することもある。

費用対効果分析

ネットの効果とともに費用（コスト）を分析する。

要するに、プログラムの目的・目標に照らして、それが適切に達成されているか、どのように達成されたのか、どの程度達成されたか、費用と比較してどうか、といった観点からプログラムの効果を評価するというものです。

（参考文献）U.S. GAO (1998) “Performance Measurement and Evaluation: Definitions and Relationships,” GAO/GGD-98-26