

# 「研究開発評価研修プログラム」テキスト

---

「文部科学省の「研究開発評価研修プログラム教材」の改定に係る調査・分析」報告書

## 目次

<b>1. はじめに</b> .....	<b>1</b>
1.1 本テキストのねらい.....	1
1.2 本テキストの構成.....	1
1.3 本テキストの使い方.....	2
<b>2. 日本における研究開発評価の枠組み</b> .....	<b>3</b>
2.1 研究開発をめぐる評価枠組みの全体像.....	3
2.2 研究開発評価.....	3
2.2.1 「国の研究開発評価に関する大綱的指針」の概要.....	3
2.2.2 「文部科学省における研究及び開発に関する評価指針」の概要.....	5
2.2.3 研究開発プログラムの評価.....	6
2.2.4 研究開発課題の評価.....	15
2.2.5 研究者等の業績の評価.....	16
2.2.6 研究開発機関等の評価.....	17
2.3 政策評価.....	17
2.3.1 政策の階層.....	18
2.3.2 3つの評価方式.....	18
2.3.3 評価の内容及び方法.....	20
2.4 行政事業レビュー.....	21
2.5 法人評価.....	22
2.5.1 独立行政法人に対する法人評価.....	22
2.5.2 国立大学法人等に対する法人評価.....	27
2.6 自己点検・評価／認証評価.....	34
2.6.1 自己点検・評価及び認証評価の導入経緯と法人評価との関係.....	34
2.6.2 自己点検・評価.....	35
2.6.3 認証評価.....	36
<b>3. 研究開発評価の基礎知識</b> .....	<b>38</b>
3.1 なぜ評価が必要なのか.....	38
3.1.1 研究開発評価の関係者とは誰か.....	38
3.1.2 環境の変化と組織経営.....	39
3.1.3 PDCA サイクルと評価.....	39
3.1.4 コンプライアンスとしての評価.....	41
3.2 評価が有用であるためには.....	42
3.2.1 評価の理念を確立する.....	42
3.2.2 評価対象を良く理解する.....	47
3.2.3 形式的な評価にしないための工夫をする.....	47
3.3 評価の留意点.....	48
3.3.1 評価の重複.....	48
3.3.2 評価の費用.....	49

3.3.3	数量的評価のデメリット	49
3.3.4	評価に対するインセンティブ	50
3.4	研究開発評価を適切に行うために学ぶべきこと	51
3.4.1	評価対象の捉え方	53
3.4.2	研究開発評価のデータと方法論	76
3.4.3	研究開発評価に関わる専門人材	95
3.5	プログラム評価	100
3.5.1	プログラム評価とは何か	100
3.5.2	評価のポイント	101
3.5.3	ロジックモデルの作成方法	104
3.5.4	評価の方法	114
3.6	プロジェクト評価	117
3.6.1	プロジェクト評価とは何か	117
3.6.2	評価のポイント	119
3.6.3	レビュー評価の手続	128
3.7	研究者等の業績の評価	134
3.7.1	研究者等の業績の評価とは何か	134
3.7.2	評価のポイント	135
3.8	研究開発機関等の評価	136
3.8.1	研究開発機関等の評価とは何か	136
3.8.2	評価のポイント	136
3.8.3	評価の内容	137
3.8.4	評価の方法	144
<b>4.</b>	<b>事例から学ぶ：より良い研究開発評価の実施に向けて</b>	<b>151</b>
4.1	科学技術イノベーション創出、課題解決のためのシステムの推進	151
4.1.1	課題の概要	151
4.1.2	科学的な挑戦性と成果の有用性を追求する競争的資金プログラム	152
4.1.3	地域課題への取組を対象とした組織内競争的研究資金プログラム	159
4.1.4	ステークホルダーとの協働によるトランスディシプリナリー研究の推進	169
4.1.5	モデル化と政策提言としての評価	186
4.2	挑戦的な研究、学際・融合領域・領域間連携研究等の推進	192
4.2.1	課題の概要	192
4.2.2	ハイリスク型研究支援のためのプログラム設計と評価	192
4.2.3	世界最先端の研究開発を推進する機関の評価	202
4.3	次代を担う若手研究者の育成・支援の推進	211
4.3.1	課題の概要	211
4.3.2	自立した若手研究者を育成するためのプログラム	211
4.3.3	特定分野における若手研究者育成のためのプログラム	214
4.4	評価の形式化・形骸化、評価負担増大に対する改善	225
4.4.1	課題の概要	225
4.4.2	自己評価の高度化のための評価	225

4.4.3 調査・分析・評価と意思決定を有機的につなげる.....	229
-----------------------------------	-----

## 1. はじめに

### 1.1 本テキストのねらい

研究開発評価には、限りある資源の中で、公平で競争的な研究環境をつくりあげるとともに、上位の政策・施策、組織の目的を達成するために、独創的で優れた課題等を発掘し、研究資金等を配分したり、施策、課題、組織の活動が適切に機能しているかを点検し、改善に結びつけたり、はたまた、組織体が、与えられた使命を実現しているかどうか説明責任を果たしたりと、多様な機能があります。

「国の研究開発評価に関する大綱的指針（以下、大綱的指針）」（平成 28 年 12 月 21 日改定、内閣総理大臣決定）は国費を用いて実施される研究開発に関する評価の全般的なガイドラインであり、この中で、「プログラム評価」を実質化することが大きな柱の一つとして掲げられました。「プログラム評価」は他の先進国では従前より一般に行われているものであり、日本でも平成 26 年に「大綱的指針」を改定する際に導入が図られましたが、その実施状況を見ると、ほとんど進展していません。その原因としては、1) 政策体系の問題（政策や組織戦略のプログラム化が図られていない）や 2) 評価体系の問題（法的に要求される評価への対応が優先される）、3) 能力及び経験の問題（意義ややり方が分からず、負担感が大きい）、といった点が指摘されています。

本テキストは、上記の 1) 及び 2) という本質的な制約がある中で、3) の課題を克服すべく、評価の初学者にも分かるよう、評価制度の全体像、評価に関する基本的な考え方や知識、各評価実施主体における具体的な手続を念頭に置いた評価のポイントについてまとめてあります。一方、真に有効な評価を行うためには、組織における政策や取組の体系を整え、しかるべき評価システムを構築していくことが求められます。したがって、評価の現場で実務を担当されている方だけではなく、こうした評価システムを構築していく権限と責任を有する組織経営の担い手向けにも役立つよう書かれてあります。

しかしながら、本テキストを読むだけで自動的に評価ができるようになるわけではありません。「評価」は、その時点での最善の知識に基づいて評価対象を深く理解するとともに、自ら立てた仮説の妥当性をしかるべき手続の下で検証し、行動の見直しにつなげる非常にクリエイティブな行為です。「評価」自体が一つの研究プロセスであると言っても過言でなく、評価を行う者の創意工夫が欠かせないのです。このように、評価の世界にはこうすればすべてがうまくいくという正解や近道はありませんが、研究プロセスですから、手がかりとなる「先行研究」があります。その意味において、本書はより確かな第一歩を踏み出すための足がかりとなるよう、先人たちの残した実績の上に立つものです。

### 1.2 本テキストの構成

本テキストは、次のような構成で成り立っています。なお、各章には、基本的な知識等について解説した本文のほか、学びを深める上で役に立つ事例や演習問題、重要な論点や先端の課題をまとめたコラムが含まれています。

まず、第 2 章では、研究開発評価に関わる制度や枠組みについて紹介するとともに、各主体にどのような評価が求められているのかを解説しています。日本における研究開発評価に関わる枠組みは多元的、多層的なものとなっており、政策評価、法人評価、自己点検・評

価といった各種の制度が、「大綱的指針」というガイドラインの下、複雑に関連しあっています。これらの全体像やそこで求められているポイントを的確に理解することが、より良い評価を行うための第一歩であると言えるでしょう。

続く第3章では、研究開発評価を行う上での基礎的な知識についてとりまとめています。ここでは、評価の基本的な考え方や「研究開発」そのものの捉え方について解説するとともに、1) プログラム評価、2) 研究開発課題の評価、3) 研究者等の業績の評価、4) 機関評価といった評価対象別に、どのような内容や方法、仕組みで評価すれば良いのかを考えます。

最後の第4章は、研究開発評価における先端の課題を実際の事例に即して考えます。具体的には、「科学技術イノベーション創出、課題解決のためのシステムの推進」、「挑戦的な研究、学際・融合領域・領域間連携研究等の推進」、「次代を担う若手研究者の育成・支援の推進」、「評価の形式化・形骸化、評価負担増大に対する改善」といった4点を取り上げています。これらは、文部科学省の所掌に係る研究開発について評価を遂行する上での基本的な考え方をまとめたガイドラインである「文部科学省における研究及び開発に関する評価指針」（平成29年4月1日文部科学大臣決定）において、特筆課題として明記されているものです。先進各国では、国として求める研究開発等の成果を生み出すためのよりよい仕組み＝プログラムのあり方をめぐって、競争が起こっています。いかによいプログラムを作れるかが競争力の源泉ともなっており、評価はそのためのツールとして欠かせないものとなっています。その意味で、本章の内容は評価の初学者や経験の浅い方にとってはアドバンストな内容であるとも言えますが、逆に言えば、政策や事業の形成・実施を担っているすべての方が関係者となる重要な論点に言及したものであるとも言えます。

### 1.3 本テキストの使い方

本テキストは、評価の実務に携わる方が自己学習を行う際の参考資料として、また、文部科学省における研究開発評価研修等の教材として使われることを想定しています。

これから評価の実務に取り組もうとされている方やこれまで自己流で評価を行って来られた方は、第2章、第3章と順に読み進め、そこで得た知識をもとに第4章の演習問題にトライしてみてください。

評価について豊富な経験のある方や体系的に学んだことのある方は、第4章を中心に、関心のあるところからご一読いただければと思います。その他の章についても、知識やスキルを整理し、再確認するのに役立つと思いますので、実務において行き詰まったり迷われた際にご参照いただければと思います。事例やコラムだけを読んで、知識を深めていただくのも良いでしょう。

本テキストが、各評価実施主体におけるプログラム等のよりよい運営と評価に向けた足がかりになれば幸いです。本書を手にとられた皆様が先駆者として評価のよりよい実践を積み上げていくことで、日本全体の底上げが図られることを心から祈念しています。

## 2. 日本における研究開発評価の枠組み

本章では、日本における研究開発に関わる評価の制度や枠組みについて紹介するとともに、各主体にどのような評価が求められているのかを解説します。

### 2.1 研究開発をめぐる評価枠組みの全体像

日本における研究開発に関わる評価の枠組みには、大きく分けて次のようなものが存在します。

- 1) 研究開発評価
- 2) 政策評価
- 3) 行政事業レビュー
- 4) 法人評価
- 5) 自己点検・評価／認証評価

これらの枠組みは独立に存在しているものではなく、相互に関連性を持ちながら、歴史的に深化してきました。評価を行う上で各主体に何が求められているのか、どのようにすれば良い評価になるのかを理解するためには、自身が直接関係している制度について熟知するだけでなく、制度間の関連性を把握する必要があります。

表 2-1 は、上記 5 つの評価の枠組みと被評価者となる組織との対応関係を簡略化し一覧にまとめたものです。本章は、読者の所属する組織に関連の深い箇所だけ読み進めることができるよう書かれてありますが、続く「2.2 研究開発評価」については全員一読することをお勧めします。

表 2-1 各評価の枠組みと被評価者となる組織との関係

	府省	国立研究開発法人等 (資金配分機関)	国立研究開発法人 (研究開発実施機関)	大学共同 利用機関	国立大学 法人	私立大、 公立大等
研究開発評価	○	○	○	○	○	○
政策評価	○	▲	▲	▲	▲	▲
行政事業レビュー	○	▲	▲	▲	▲	▲
法人評価		○	○	○	○	
自己点検・評価/ 認証評価					○	○

注) ▲は間接的に被評価者となりうることを示す。

### 2.2 研究開発評価

#### 2.2.1 「国の研究開発評価に関する大綱的指針」の概要

日本では、科学技術創造立国の実現を目指して「科学技術基本法」(平成 7 年)が制定されており、また、それを具体的に推進していくための国家戦略として、同法に基づいて 5 年ごとに「科学技術基本計画」(以下、「基本計画」。現行のものは平成 28~平成 32 年度を対象とする第 5 期)が策定されています。平成 8 年に策定された第 1 期基本計画では、「研究開発機関及び研究開発課題について、評価の在り方を抜本的に見直し、適切な評価の仕組

みを整備し、厳正な評価を実施」することが求められたことから、「国の研究開発全般に共通する評価の実施方法の在り方についての大綱的指針」（平成 9 年 8 月 7 日内閣総理大臣決定）を策定し、研究開発評価の導入と定着化を進めてきました。平成 13 年 11 月には「国の研究開発評価に関する大綱的指針」（以下、「大綱的指針」）と名称を改め、評価の公正さと透明性の確保、評価結果の資源配分への反映、評価に必要な資源の確保と評価体制の整備などを追加して、厳正な評価を推進する内容に改定されました。その後、基本計画の改定等にあわせて、国際水準の信頼できる評価の実施や研究開発プログラム評価の導入など内容を変更してきた、という経緯があります。

日本における公的資金に基づく研究開発活動の運営や実施に関わる評価は、この大綱的指針によって規定されており、現行の大綱的指針は第 6 版（平成 28 年 12 月 21 日内閣総理大臣決定）になります。第 6 版は、第 5 期科学技術基本計画を受けて策定されたものであり、1) 実効性のある研究開発プログラム評価のさらなる推進、2) 挑戦的（チャレンジング）な研究開発等の促進、3) 研究開発評価に係る負担の軽減といった観点が特に重視されています。

さて、以下では、国の研究開発全般に共通する評価のガイドラインである大綱的指針の中身についてみていくことにします。まず、大綱的指針においては、その範囲について、次のように記載されています。

研究開発の範囲は、国費を用いて実施される研究開発とする。具体的には、各府省、国立研究開発法人等及び大学等が自ら実施又は推進する研究開発が対象となる。また、民間機関や公設試験研究機関等で国費の支出を受けて実施される研究開発、国費により海外で実施される研究開発等も対象とする。

つまり、「国費を用いて実施される研究開発」が関わっているものであれば、すべてが大綱的指針の対象となります。詳しくは後述しますが、これには、機関レベルで対応するもの、研究者個人やチームレベルで対応するものの両方が含まれます。

このように、大綱的指針は、国費に関わるすべて研究開発関連活動の評価を対象とする最も基本となるガイドラインですが、その記述内容は最小限のものにとどまっています。これは、「研究開発が関連する政策・施策等の目的や、その目的達成のために実施される研究開発の特性等に応じて、評価の実施主体が独自の工夫に基づき評価を実施できるようにする」ためであり、大綱的指針では、「あらゆる関係者が高い当事者意識を持ちつつ自らの責任の下で自発的に対応すること」を求めています。なお、ここでいう「政策・施策等」には、府省の行うものだけでなく、広義には国立研究開発法人等や大学等が自らの戦略（中（長）期目標・計画など）の下で実施するものも含まれます。

このような問題意識の下、評価の実施主体である各府省は大綱的指針や後述する「行政機関が行う政策の評価に関する法律」（平成 13 年法律第 86 号）（以下、「政策評価法」）等の関係法令に沿って、評価対象、評価の目的及び評価結果の取扱い、評価時期、評価方法など評価の実施に関する事項について、研究開発評価の指針を自ら策定することになっています。文部科学省については、「文部科学省における研究及び開発に関する評価指針」（平成 29 年 4 月 1 日文部科学大臣決定）がこれに該当します。

また、各府省が所管する国立研究開発法人等や大学等の研究開発機関、第三者評価機関<sup>1</sup>においても、大綱的指針、各府省の指針及び関係法令に沿って、同様な事項について、明確なルールを定めることになっています。

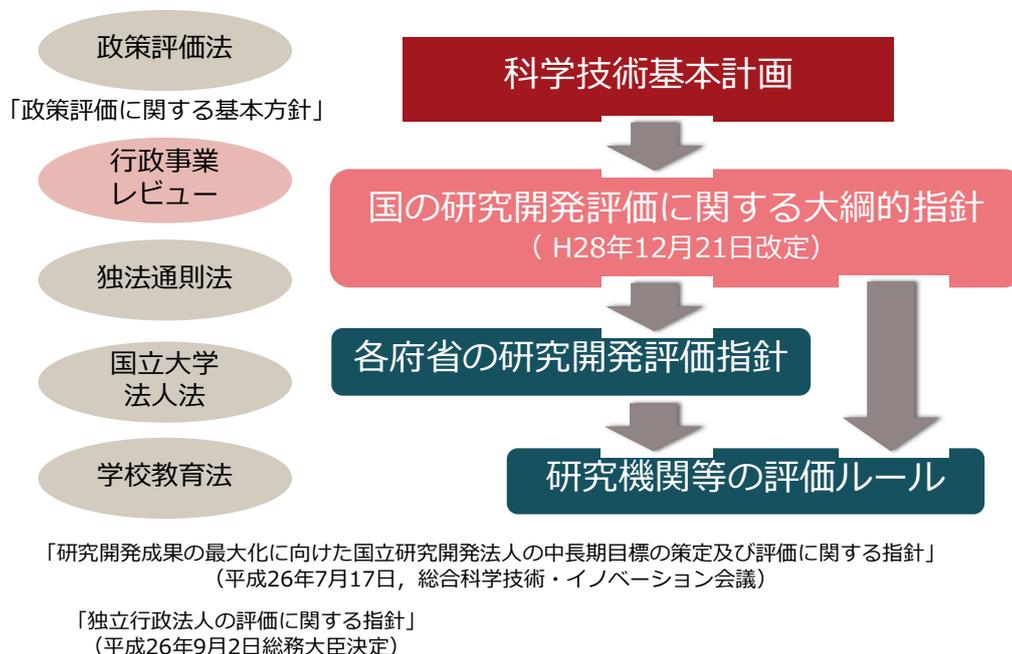


図 2-1 大綱的指針とその他の関連法等との関係

出典: 未来工学研究所作成

大綱的指針が対象とする研究開発評価には、大きく 1) 研究開発プログラムの評価、2) 研究開発課題の評価、3) 研究者等の業績の評価、4) 研究開発機関等の評価の 4 種類があります。これらがどのようなものかについては後述しますが、府省であれ、国立研究開発法人等であれ、大学等であれ、国費を用いた研究開発を支援、実施する機関は、4 種類の評価すべてに対し、「**高い当事者意識を持ちつつ自らの責任の下で自発的に対応すること**」が求められていると言えます。

### 2.2.2 「文部科学省における研究及び開発に関する評価指針」の概要

前述のように、文部科学省においても、大綱的指針に基づいて、「文部科学省における研究及び開発に関する評価指針」（以下、「文科省評価指針」）が策定されています。これは、「文部科学省の所掌に係る研究開発について評価を遂行する上での基本的な考え方をまとめたガイドライン」であり、平成 29 年 4 月 1 日に文部科学大臣決定されたものが最新版になります。

文科省評価指針の構成としては、科学技術イノベーション創出の促進に関する重要性の増大など、社会や時代が抱える喫緊の諸課題等との関係で特に研究開発評価が積極的に対応し

<sup>1</sup> 評価の対象となる研究開発の実施・推進主体とは別の独立した機関であり、総合科学技術・イノベーション会議、国立大学法人評価委員会、大学改革支援・学位授与機構等が念頭におかれています。

ていくべきと考えられる課題について、第1部において特筆するとともに、第2部において研究開発評価に係る「基本的事項」、「対象別事項」等について整理しています。また、巻末において、指針の中で用いられている用語の概念等の整理や指針の記述内容の理解を深める参考となる関連情報の整理を行っています。

文科省評価指針においては、文部科学省内部部局について、「本指針に基づき、実施要領を策定するなど所要の評価の枠組みを整備し、自らの研究開発に関する評価を行うこととする」とされているほか、大学等や文部科学省所管の研究開発法人等については、「本指針を参考にしつつ、自らがその特性や研究開発の目的・手法・性質等に応じて多様で柔軟な評価システムを構築し、それぞれ適切な方法により進めることが期待される」としています。

また、関連法令等との関係についても言及されており、「政策評価法」に基づく評価のうち、研究開発を対象とする政策評価を実施するに当たっては、「大綱的指針及び本指針に基づき行うものとする」とされているほか、研究開発機関等の評価のうち、研究開発法人等については、「「独立行政法人通則法」（平成11年法律第103号）及び「特定国立研究開発法人による研究開発等の促進に関する特別措置法」（平成28年法律第43号）に基づく評価」を、国立大学法人及び大学共同利用機関法人については、その研究活動の特殊性に鑑みて、「国立大学法人法」（平成15年法律第112号）に基づく評価が行われるが、それに当たっては、本指針を参考とすることが期待される、とされています。

以下では、大綱的指針及び文科省評価指針に基づいて、対象別に解説します。

### 2.2.3 研究開発プログラムの評価

プログラム評価は、研究開発に限らず、先進各国では最も基本的な評価の単位として、従前より一般に行われてきました。一方、日本においては大綱的指針の第5版（平成24年12月6日内閣総理大臣決定）から導入され、前述のように、大綱的指針の第6版でも「実効性のある研究開発プログラム評価のさらなる推進」が強調されていますが、これはプログラム評価が思うように進んでいなかったという状況を反映しています。その理由は、「1. はじめに」でも触れたように、1) 政策体系の問題や2) 評価体系の問題、3) 能力及び経験の問題といった点が指摘できます。

以下ではまず、こうした問題がどのようなものなのか、また、それらに対し、大綱的指針や文科省評価指針ではどのように対応しようとしているのかをみていきます。

#### (1) プログラム評価が根付かない理由とそれらへの対応

##### 1) 政策体系の問題

まず、政策体系の問題ですが、評価の対象とする「政策」がどのような目的の下でどのような手段を用いるものかという対応関係が明らかでないと、実効的な評価は行えません。政策評価法に基づく政策評価の円滑かつ効率的な実施のための標準的な指針を示した「政策評価の実施に関するガイドライン」（平成17年12月16日政策評価各府省連絡会議了承、平成24年3月27日一部改正）においても、「政策評価の体系的かつ合理的で的確な実施を確保するため、政策体系をあらかじめ明らかにすること」が基本とされ、その区分の仕方は様々

であるとしながら、「政策（狭義）-施策-事務事業」などの政策体系をあらかじめ明示した上で評価を実施することが必要」であることが述べられています。しかしながら、現実の政策体系は、最上位の政策から目的-手段間の関係に基づいて論理的にブレイクダウンされたものというよりも、本来手段である既存の政策（事務事業等）を目的の類似性などをもとに後から束ね、それを上位の政策と紐付けたと思われるものもあり、そうしたものについては目的-手段間の関係が明瞭でないとの意見もあります。また、「政策（狭義）」、「施策」、「事務事業」の区分がそもそも相対的な概念であり、これらと「プログラム」をどのように関連づけて考えてよいか分かりにくいといった点も指摘できます。さらには、現実の政策においては、明らかにプログラムとして扱われるべきものがプロジェクトと呼ばれているなど、言葉をめぐる混乱もみられます。これは行政機関に限らず、国立研究開発法人等や大学等においても同様の状況が少なからずみられます。

こうした状況を改善するため、大綱的指針第 6 版では、「研究開発プログラム」を次のように定義しました。

『研究開発プログラム』とは、研究開発が関連する政策・施策等の目的（ビジョン；何のためにやるのか）に対し、それを実現するための活動のまとまりとして位置付けられる。

すなわち、研究開発が関連する政策・施策等も『研究開発プログラム』ということができるとともに、政策・施策より下位の階層における事務事業等も『研究開発プログラム』になる場合がある。また、競争的資金制度等の研究資金制度なども『研究開発プログラム』の 1 つである。

この『研究開発プログラム』の範囲を組織の観点から見ると、課内にとどまるもの、課をまたがるもの、局をまたがるもの、府省をまたがるもの、独立行政法人の分野・領域単位のもの、複数の独立行政法人にまたがるもの等、運用する組織、機関のガバナンスの下で様々な範囲の『研究開発プログラム』が設定可能である。

このように、「研究開発プログラム」は特定の階層区分や組織区分に紐づけられるものではなく、**目的-手段の論理的な体系群**であることが明らかにされました。

## 2) 評価体系の問題

2 点目は、評価体系の問題です。詳細は後述しますが、府省は政策評価法に加え、予算と連動する行政事業レビュー、国立研究開発法人等は独立行政法人通則法に基づく法人評価、大学共同利用機関や国立大学等は国立大学法人法に基づく法人評価といった法的に要求される評価や政権の要請で実施される評価があり、これらへの対応を優先せざるを得ません。リソースに限りがある中、これらに加えて別途「プログラム評価」を行うとなると、評価の実施主体に過剰な負担がかかるのではないかという懸念が生じるのはある意味当然の反応であると言えます。大綱的指針はガイドラインですので、「プログラム評価」をたとえ行わなかったとしても罰則等はありません。そのため、評価を実施するインセンティブも一見ないように思えます。

こうした状況を踏まえ、大綱的指針では、次のように「他の評価体系との整合」を図ることの重要性が強調されています。

・・・本指針は、政策評価に求められている諸要素を踏まえ、さらに、研究開発の特性を考慮したうえでの留意事項を示すものであり、本指針による評価の実施にあたっては、政策評価法に基づく政策評価と整合するよう取り組むこととする。また、研究開発機関等の評価のうち、国立研究開発法人等については「独立行政法人通則法」(平成11年法律第103号)に基づく評価、さらに国立大学法人及び大学共同利用機関法人については「国立大学法人法」(平成15年法律第112号)に基づく評価と整合するよう取り組むこととする。

・・・研究開発が関連する政策・施策等も『研究開発プログラム』ということができること、また、研究開発評価は政策評価等と整合するよう取り組むこととされていることを踏まえ、『研究開発プログラムの評価』の充実を図っていく手段として、例えば、同一の評価対象に対して、『研究開発プログラムの評価』を政策評価とは別に実施する必要が生じないように取り組むなどにより関係者の意識向上と評価の効率化を図り、その後、段階的に『研究開発プログラムの評価』の水準の向上を図ることが考えられる。

このことから、大綱的指針においては、既存の評価体系における評価を実施する際に、**研究開発の特性を考慮することに加え、そこにプログラム評価の概念を持ち込むことで既存の評価をアップグレードすること**を求めていると考えることができます。これを実現するためには、「全ての関係者が、今以上に評価の本質を理解し、その上で評価の意義に沿った行動をとることが必要」不可欠であり、「評価の客観性や正当性を確保するという前提の下、**評価の実施主体による自己評価**」を**評価の基盤として取り組んでいくことが必要**です。

また、国費を用いて実施される研究開発は、それらを実施又は推進する主体の面から見ても、評価の対象となる研究開発の面から見ても、階層構造となっており、「研究開発評価は、このような階層構造の下で各々の階層において重層的に実施されることから、同一の研究開発が複数の階層における評価の対象とされることが多い」のも事実です。いわゆる「評価疲れ」を引き起こさないためには、「同じような評価が繰り返し行われないこと、形骸化した評価を実施しないこと」、つまり、下位の階層などですでに実施された評価の結果を上位の階層で行う評価に有効活用することや、評価対象の改善など意思決定にきちんと反映させることが重要になります。これを実現するためにも、**評価計画をあらかじめしっかりと立て、結果を含めて関係者間で共有すること、そして、「評価の実施主体による自己評価の質を向上させることにより、その結果をその後の上位階層での評価において、そのまま活用できるようなものとする**こと」が求められます。

### 3) 能力及び経験の問題

「評価の実施主体による自己評価の質を向上させること」が重要であることを先に述べましたが、日本において、プログラム評価の経験が全般的に不足している中でこれを行っていくのは困難です。ただでさえ、「研究開発評価を支えるための人や予算」が不足している中、プログラム評価を行うことにどのようなメリットがあるのかと疑問に思われる方も多いでしょう。

本テキストは、ある意味こうした状況を克服するためにとりまとめられたものですが、大

網的指針においても「本指針が各府省や研究開発機関等、評価に関係するあらゆる関係者に浸透し、真に実効性をもつものとするためには、改善の努力を継続していくことが極めて重要」であり、「できる部分から段階的に運用し、評価方法の改善のための調査研究や評価の実施状況等のフォローアップを継続して行う」と述べられています。

また、文科省評価指針においても、「いかなる評価システムが構築されても実際に評価を行うにふさわしい人材が乏しければ適切に機能しないことから明らかなように、評価に責任と権限を持つことができる、評価に関わる資質能力を備えた人材を育成していくとともに、当該人材が活躍できる環境やキャリアパスを整備していくことも重要な課題」であると指摘されています。

先進各国では、国として求める研究開発等の成果を生み出すためのよりよい仕組み＝プログラムのあり方をめぐって、競争が起こっています。いかによいプログラムを作れるかが競争力の源泉ともなっており、評価はそのためのツールとして欠かせないものとなっています。経験の乏しい日本において、あらゆる関係者が「習うより慣れよ」の精神でまずはプログラム評価に取り組んでいくことが求められているといえますが、海外の先行する取り組みやこれまで蓄積された評価研究の成果等を参照することで、より質の高い試行が可能になっていくでしょう。

## (2) 研究開発プログラムとプログラム評価

さて、ここまで「プログラム評価」とは何かについて言及しないまま、説明してきました。また、評価の対象である「研究開発プログラム」についても「目的-手段の論理的な体系群」であるということを示しただけで、詳細については触れておりません。ここでは、その特性について、大綱的指針や文科省評価指針を参考にしながらもう少し明確にしていきたいと思えます。

まず、大綱的指針では、次のように言及されています。

研究開発プログラムでは、研究開発を実施する主体ではなく政策立案者や推進する主体等に第一義的な「責任の所在」がある。政策・施策等を立案・推進する側とその下で研究開発を実施する側との役割分担と責任の所在を明確化し、概念としては、政策立案者や推進する主体等の行動及びその結果について評価を行うのが、『研究開発プログラムの評価』である。・・・プログラムの改善や次のプログラム立案のための示唆を得る。

・・・国費を用いてイノベーションを生み出すためには、あるべき社会の姿の実現を政策・施策等の目的として、具体的な政策・施策等の目標を設定し、それに必要な研究開発課題等の活動を組み合わせて実行することとなる。

このとき、これらの活動のまとめりとして構成した『プログラム』の単位で研究開発を推進し、『プログラム』を推進する主体の行動及びその結果を評価していくことが重要である・・・。

また、文科省評価指針では、次のように定義されています。

研究開発プログラムとは、複数の研究開発課題を運営する施策や競争的資金制度等、研究開発に係る政策上の特定の目的や目標の実現を目指して、推進方針や戦略・計画・実施手段等の体系が整備され、それに応じて推進されるものをいう。

研究開発プログラムの評価は、・・・このような施策や制度等を対象として、目標の設定された研究開発プログラムごとに評価を実施することにより、実施の可否を判断するとともに、研究開発の質の向上や運営改善、計画の見直し等につなげることを目的とする。

これらを総合すると、「研究開発プログラム」とは、研究開発に関わる政策上の特定の目的や目標を実現するための「目的-手段の論理的な体系群」であり、「**研究開発を実施する主体ではなく政策立案者や推進する主体に第一義的な「責任の所在」がある**」ものである、ということになるでしょう。そもそも評価は「**何らかの意思決定（資源配分、改善・質の向上、進捗度の点検、説明責任等）を行う目的のために実施される手段**」（文科省評価指針）ですが、「研究開発プログラム評価」は、その意味で「**政策立案者や推進する主体等の行動及びその結果について評価を行う**」ものであり、プログラムの「**実施の可否を判断するとともに、研究開発の質の向上や運営改善、計画の見直し等につなげること**」や「**プログラムの改善や次のプログラム立案のための示唆を得る**」ために行うものである、といえます。

また、「研究開発プログラム評価」についてより深く理解するためには、「研究開発」に関わるからこそその特性について考える必要があります。研究開発に関わる政策は、各国においても、研究政策、研究開発政策、科学技術イノベーション政策など様々な呼称が用いられるなど「各国の歴史的背景の違いや時代の変化とともにその内容は変遷してきて」おり、「明確で、権威のある定義はないに等しい」と言える状態ですが、次のような共通した特徴があります（小林 2011<sup>2</sup>，田原 2014<sup>3</sup>）。

まず、科学技術が政策の目的であると同時に、政策の手段でもある、という二面性です。小林は、1971年に発表されたOECDによる報告書（通称、ブルックス報告）を引き、前者を「科学のための政策(policy for science)」、後者を「政策のための科学(science for policy)」と呼びましたが、この両者は明確には分離できないという性質を持っています。なぜなら、「長い歴史の中で蓄積された科学技術の知識であっても、そのままでは個別の政策課題に対して適切な知見を提供できるとは限らず、新規の研究開発が必要になる場合も少なくない」からであり、その場合、個別政策分野の問題解決のために動員される科学技術能力の結集施策 (science for policy) は、「科学技術の振興施策 (policy for science) に限りなく近いものとなる」からです。その意味で、「科学技術の振興政策と個別政策の中で取り込まれる科学技術能力の結集政策の両面を持ちつつも、それらを別の政策としてではなく、包括的に対象とする政策領域として成立してきたのが、科学技術政策」であるといえます。また、科学者コミュニティは、科学技術政策において政策的支援を受ける対象である一方で、その資源配分の決定等にも専門家として関与しています。つまり、科学技術政策は、「科学技術が政策を支える専門性であると同時に政策の対象であるという反射的（自己言及的）」な特性を

<sup>2</sup> 小林信一、「科学技術政策とは何か」『科学技術政策の国際的な動向 [本編]』（国立国会図書館調査資料 2010-3），pp.7-34，国立国会図書館，2011。

<sup>3</sup> 田原敬一郎，「公共政策学」，山下晋司編『公共人類学』，東京大学出版会，2014。

持っており、このことが科学技術政策を「他の政策にはない際立った」ものにしており、同時に、「複雑性の源泉となっている」といえます（小林 2011）。

さらに、科学技術政策は、上記のような二面性だけではなく、科学技術がもたらす負の側面への対応、という性格も持ち合わせています。つまり、科学技術が問題それ自体となる場合であり、科学技術の専門家は問題解決の担い手であると同時に、問題を引き起こす当事者ともなりえます。専門家に対する社会の信頼はこのような中で揺れ動き、そのことが科学振興に対する投資や規制のあり方にも大きな影響を及ぼすことにもなっています（田原 2014）。

こうした「研究開発」「科学技術」に関わる政策の特徴は、いわゆる政府による「公共政策」だけにみられるものではなく、研究開発法人や大学等の研究開発を扱う主体が行う事業等についても当てはまります。

こうした特徴を踏まえた上で、「研究開発に係る政策上の特定の目的や目標」を類型化すると、原理的には、「科学的価値」、「経済的価値」、「社会的価値」といった3つの価値のいずれかを実現するものとして分類することができます。

表 2-2 実現を目指す価値別に見た研究開発プログラムの特徴

価値区分	成果の受け手	価値発現のメカニズム
科学的価値 (ディシプリン志向)	研究者(ピア)。知識生産者と同じコミュニティ	ジャーナルや高等教育システムを通じて
経済的価値 (ミッション志向)	企業、消費者等。知識生産者と異なるコミュニティ	市場を通じて
社会的価値 (ミッション志向)	利害関係者、政策決定者、市民等。知識生産者と異なるコミュニティ	公的秩序もしくは市民的秩序を通じて(社会実装・社会的受容、政策過程・政治的受容)

出典:「海外の府省及び資金配分機関等における研究開発プログラム及びプログラム評価に関する調査・分析」調査報告書(平成 25 年度文部科学省委託調査)

これらの3つは、価値実現のための手段として研究開発や専門知が関わる点では共通していますが、成果の受け手や研究開発と価値実現とをつなぐメカニズムの点で異なっています。近年では、ステークホルダー協働で行うトランスディシプリナリー型の研究や多様な主体が関わるオープンイノベーション型の研究開発なども行われるようになっており、単純に知識の生産者と利用者を分けることはできませんが、いずれにせよ、実現を目指す価値が異なればその実現のための手段や道筋も異なりますし、評価の視点も変わってきます。

「研究開発プログラム評価」を行うにあたっては、こうした研究開発が関わるからこそその特性を踏まえた上で、評価対象とする研究開発プログラムがどのような価値の実現を目指すものかを明確にすることが最初の一步として非常に重要であると言えます。

なお、上述の文科省評価指針においても示唆されているように、研究開発プログラムは競争的資金制度に限ったものではありません。たとえば、アメリカの連邦政府関連機関では、運営費交付金のようなブロックファンドや規制などを含まずすべての政策がプログラム化されており、プログラムを基本単位とした評価が伝統的に実施されています<sup>4</sup>。評価を行う際

<sup>4</sup> 米国の大統領府行政管理予算局 (OMB) がまとめた「プログラム評価評定ツール (PART) を完了するためのガイダンス」より。なお、プログラム評価評定ツール (The Program Assessment Rating Tool: PART) はブッシュ政権時に OMB が考案し導入したものであり、プログラムごとにその実施状況や成果を評価する仕組みです。オバマ政権やトランプ政権では用いられていませんが、その内容は普遍的な要素を持っています。

には、上記のような実現を目指す価値の違いだけではなく、こうした手段の性質の違いについても理解しておく必要があります。

表 2-3 アメリカ PART で用いられていたプログラムの類型

類型	概要
①直接的連邦プログラム (Direct Federal Programs)	第一義的に、連邦政府の職員がサービスを提供するもの
②競争的助成プログラム (Competitive Grant Programs)	競争的なプロセスを経て、州、地方及び部族の政府、組織、個人その他の主体に資金を提供するもの
③ブロック／定型助成プログラム (Block/Formula Grant Programs)	定型のもしくは定額交付金として、州、地方及び部族の政府、組織、個人その他の主体に資金を提供するもの
④規制型プログラム (Regulatory-Based Programs)	法律や政策を導入、解釈、規定したり、手続や実施要件を説明するルール形成を通じて、そのミッションを達成しようとするもの
⑤固定資産・サービス調達プログラム (Capital Assets and Service Acquisition Programs)	(土地、建造物、設備及び知財等の)固定資産の開発や取得、(メンテナンスや情報技術等の)サービスの購入を通じて、その目的を達成しようとするもの
⑥クレジット・プログラム (Credit Programs)	融資や借入助成及び直接控除を通じて支援を提供するもの
⑦研究開発プログラム (R&D Programs)	知識の創造、もしくはシステムや方法、構成要素、技術の創造に対する活用に焦点をあてたもの

※「研究開発プログラム」かつ「競争的助成プログラム」といった複数類型をまたぐものもある。

出典：Office of Management and Budgeting (OMB), Guidance for Completing the Program Assessment Rating Tool (PART), March 2005.をもとに未来工学研究所作成。

### (3) プログラム評価において重要な「道筋」

先に、「研究開発プログラム」とは「目的-手段の論理的な体系群」であると述べましたが、これを具体化したものが「道筋」です。大綱的指針では、この「道筋」について次のように定義しています。

・・・『道筋』とは、政策・施策等の目的に対し、現状がどうなっているか、政策・施策等の目的と現状のギャップを埋めるためにどんな活動をどの順番で行うか、成果の受け手側で発現することが期待される効果・効用等を時間軸に沿って描いたものである。具体的には、時間軸に沿って、目的達成に影響を与える外部要因や受益者、期待しているアウトカム、政策・施策等の目標の達成度を測るための指標（アウトカム指標）、プログラムを構成する個々の活動とそれぞれの目的及び目標（アウトカム指標、アウトプット指標）、個々の活動が政策・施策等の目的に与える影響、研究開発プログラム及び研究開発プログラムを構成する個々の活動の実施者の役割・権限と責任の所在等を仮説として示したものである。

大綱的指針においては、評価を行うにあたり、そもそも**研究開発プログラムの立案段階**において、「研究開発が関連する政策・施策等の目的を達成するために、研究開発プログラムの実施者の手を離れた後で成果の享受者である成果の受け手に対して、何を、いつまでに、どの程度届けるか」といった具体的で実現可能な目標（アウトプット目標）と、成果の受け手

が行う活動及びその効果・効用として現れる価値（アウトカム目標）を、時間軸に沿った『道筋』として示すことにより、誰の責任で、何をどのように実施するのかを明らかにしておくことが重要である」としています。この「道筋」を描くためのツールとしてはロジックモデルがあり、様々な表現形態がありますが、標準的なものを図 2-2 に示します。なお、ロジックモデルの詳細については続く第 3 章で詳細に言及します。

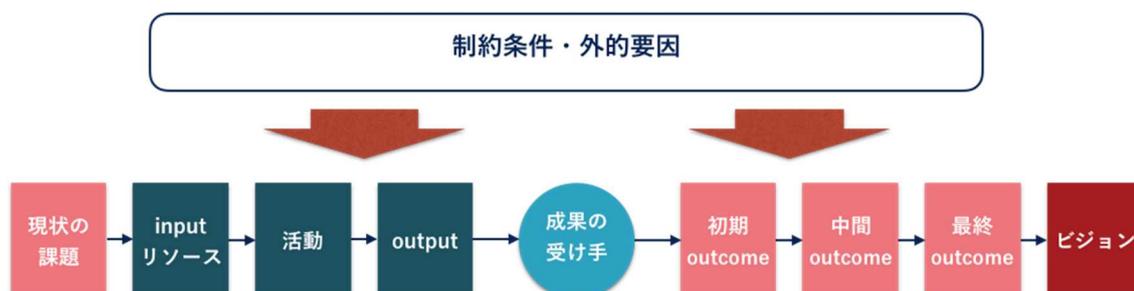


図 2-2 一般的なロジックモデル

(出典) 林隆之氏資料をもとに作成

さて、こうした観点からすると、「道筋」は事業を行う上でのある種の仮説であり、その妥当性を検証し、その実施の可否を含めて、事業をよりよい方向へと導くための手段が「プログラム評価」である、ということがいえます。つまり、研究開発を行うにあたって仮説検証プロセスが重視されるように、研究開発を推進するための仕組み＝プログラムにおいても、仮説検証プロセス＝プログラム評価を大切にしましょう、というメッセージに他なりません。当然仮説ですから、「試行を重ね、段階的に充実、見直しを図っていくべき」ものでもあります。大綱的指針では、この道筋を「作り上げるプロセスが重要であり、関係者間で『道筋』の内容の妥当性について議論を積み重ね、その内容の充実を図るとともに、全ての関係者間で『道筋』の内容についてコンセンサスを得た上で、当該研究開発プログラムを始めることが重要である」とも述べられています<sup>5</sup>。

なお、上記の大綱的指針の記述の中に、「アウトプット」、「アウトカム」という用語がでてきます。大綱的指針では、前者について「研究開発に係る活動の成果物。目的達成に向けた活動の水準を表す」、後者について「研究開発に係る活動自体やそのアウトプットによって、その受け手に、研究開発を実施または推進する主体が意図する範囲でもたらされる効果・効用」とそれぞれ定義していますが、この 2 つは重要なキーワードであり、続く第 3 章で詳細にとりあげることにします。

#### コラム 1：プログラムとプロジェクトの違い

##### プログラムおよびプロジェクトの辞書的定義

大綱的指針においては、「道筋」のような論理体系を明示化するという考え方について、「個別の研究開発課題を立案する際にも応用できるものである」としています。では、プログラムとプロジェクトとは、どのような点が本質的に異なるのでしょうか。

<sup>5</sup> 大綱的指針においては、こうした考え方について、「個別の研究開発課題を立案する際にも応用できるものである」としています。

まず、国際的に標準とされているプロジェクト・マネジメントの知識体系（ガイド、手法、メソドロジー、ベストプラクティス）をまとめた「PMBOK ガイド（第5版）」を紐解くと、プロジェクトとは「独自のプロダクト、サービス、所産を創造するために実施する有期性のある業務」であり、プログラムは「複数のプロジェクト」である、とされています。

さらに、英英辞典（American Heritage 4th ed.）をひくと、プロジェクトとは「協調的な努力を必要とする請負ごと」、プログラムは「通常、特定の社会的ニーズを満たすようデザインされた複数のサービス、機会もしくはプロジェクトからなる一つのシステム」と出てきます。

これら2つの定義を勘案すると、プロジェクトは「（与えられた資源制約の中で）協調的な努力によって期限内に活動を遂行し、目標とする何かを生み出す行為」と定義することが可能ですが、それを複数持つプログラムとは何か、そのマネジメントにおいてどのような違いが生まれるかを考える必要があります。

#### 鍵となる概念としての「システム」

ここで、両者を区分する上で鍵となる概念が「システム」です。システムとは、「構成要素が互いに関連しあうことによって機能する、複雑な集合体」（木嶋ら 2007）であり、1) 創発特性、2) 階層構造、3) コミュニケーションとコントロールという3つの特性を持ちます（Checkland 1981）。具体的には次のようなものです。

- 個同士の関係性によって生じ、かつ個に還元しえない全体の性質＝創発特性を持つ。階層が異なると、下位の階層にはなかった新たな性質・特徴がそこから現れる。
- 個々のプロジェクトもいくつかのサブプロジェクトからなる一つのシステムとして捉えることが可能。ただし、プロジェクトに還元できない全体としての性質をプログラムは持つ。
- システムを取り巻く外部環境に適応していくために（システムの中と外とを分けるものを「システム境界」と呼ぶ）、システムを構成する個々の要素間で、階層間で、そして、システムと外部環境との間で、多様な形の相互作用＝コミュニケーションを行う。
- システムの目的から外れる行動があった場合に、「情報」としてそれを取りだし、この情報をもとに目的に向かわせる修正行動＝コントロールをとる。

#### プログラムの要件

こうしたシステムとしての特性を踏まえると、プログラムとは次のような要件を満たすものである、ということができます。

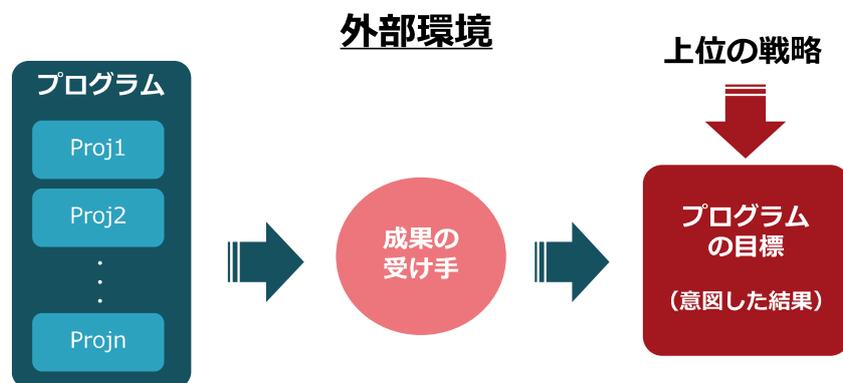


図 2-3 プログラムの概念図

出典:未来工学研究所作成

- 上位の戦略等から与えられる独自の目的・目標を持ち、それを達成するためのプロジェクトや全体性を確保するための諸活動で構成。
- プログラムの外にいる特定の社会的ニーズの持ち主に成果を届け、ニーズを充足させるまでが使命。
- そのための「コミュニケーションとコントロール」＝マネジメントの仕組みを持つ。

こうしたマネジメントを行う上で強力な武器となるものが評価であると言えるでしょう。

#### 参考文献

木嶋恭一ら編著、『ホリスティック・クリエイティブ・マネジメント』，丸善，2007年。

Checkland, P.B. *Systems Thinking, Systems Practice*, 1981. (チェックランド『新しいシステムアプローチ』オーム社)

## 2.2.4 研究開発課題の評価

研究開発課題とは、文科省評価指針によれば、「政策や研究開発プログラムの下で個別の具体的な研究開発活動が実施される単位」です。大綱的指針では、「比較的規模の大きい研究開発課題や複数の研究開発課題からなるもの等については、「プロジェクト」と称される場合もある」とされていますが、国際的には **program**（もしくは **programme**）に対比される概念として、規模の大小等に関わらずすべて **project** と呼ばれています。本テキストでは、こうした国際的な慣例に倣い、「研究開発課題」と「プロジェクト」を同義のものとして扱います。

この研究開発課題について、文科省評価指針では次の3つのタイプに分けています。

- 1) 競争的資金による研究開発課題:公募により複数の候補の中から優れたものが競争的に選択されて実施されるもの
- 2) 重点的資金による研究開発課題:国が定めた明確な目的や目標に沿って重点的に推進されるもの
- 3) 基盤的資金による研究開発課題:研究開発機関等に運営費交付金等として配分された資金により実施されるもの

まず、1) 競争的資金による研究開発課題とは、競争的資金（制度）という特定の政策目的を持つプログラムの下で実施されるものであり、研究者や研究開発実施機関は、この目的の範囲内で独自の発想に基づいてボトムアップで提案を行います。そして、プログラム運営者は、提案された複数の候補の中から、プログラムの目的に合致する優れたプロジェクトを選定します。これはプロジェクトの事前評価にあたるもので、採択審査と呼ばれることもあります。こうして採択されたプロジェクトに対し、途上段階で評価を行ったり、終了時評価が行われたりします。プログラム評価の一環として、それを構成するプロジェクトの追跡調査が行われることもあります。

2) 重点的資金による研究開発課題とは、国が定めた政策や研究開発プログラムの目的・目標を達成するために、基本的にはトップダウンでプロジェクトを決定し、それを最も効率よく着実に実施してくれることが期待できる研究開発実施主体を選定するものです。したが

って、プロジェクトの選定、すなわち事前評価にあたっては、独創性というよりも、研究者や機関の持つ研究開発の遂行能力が問われます。途上段階においては、計画通りに進捗しているかを確認したり、それに基づいて改善を促すなどの目的で評価を実施します。

3) 基盤的資金による研究開発課題とは、大学等においては「競争的資金の獲得に至るまでの構想段階の研究を保障し日常的な教育研究活動を支えるとともに、大学附置研究所、研究センターの整備や特定の大学共同利用機関等が中心となる特殊大型施設・設備を要する大規模研究の推進に大きな役割を果たすもの」です。たとえば、機関の長の裁量等で組織戦略に基づいて実施する組織内の研究費配分がこれに含まれます。これを組織内の公募で行うのであれば競争的資金におけるプロジェクト評価に準じることになりますし、トップダウンで行うのであれば重点的資金によるプロジェクトと同様の考え方で評価を実施することになるでしょう。

ただし、文科省評価指針においても指摘されていますが、こうした区分は排他的なものではなく、上記のように、複数の区分に該当することもありえます。そのため、これらの3つの区分はあまり気にせず、プログラムの目的に応じて、最適な研究開発課題（プロジェクト）の評価システムを構築していくことが求められます。

## 2.2.5 研究者等の業績の評価

研究者等の業績の評価は、当該研究者等を雇用する機関が評価実施主体となって行うものであり、「(研究者自身の)自己点検による意識改革、研究の質の向上、教育の質の向上、社会貢献の推進、組織運営の評価・改善のための資料収集、社会に対する説明責任等」を目的として実施されるものです（文科省評価指針）。

また、大綱的指針にも明記されているように、研究者等の業績の評価は科学技術システム改革の一環として位置づけられており、研究者の処遇など人事考課と密接に結びつくものであるため、「能力や業績の公正な評価の上、優れた努力に積極的に報いること等による公正で透明性の高い人事システムの徹底が重要」であると言えるでしょう。

こうした評価を実効的なものにするためには、「(研究者が)所属する機関の長が当該機関の設置目的等に照らして適切かつ効率的な評価のためのルールを整備」することや、評価結果について「インセンティブとなるよう個人の処遇や研究費の配分等に反映させる」ことが必要であり、「研究者を萎縮させず果敢な挑戦を促す」などの工夫を行う必要があります。その際、研究開発の実績だけではなく、社会貢献や教育、機関の運営管理業務に対する寄与など多面的にみていくことが求められる、と大綱的指針では述べられています。文科省指針では、「個人の能力が最大限に発揮されるとともに、組織力の向上も目指した評価となるように評価される領域（「研究」「人材育成」「社会貢献」「運営管理」等）の比重を適宜変えるなど「一律的な評価を避ける」ことや「評価項目全体を平均的に判断するばかりではなく、場合によっては、優れている点を積極的に取り上げる」ことの必要性が指摘されているほか、各研究開発機関等において「公正でかつ透明性の高い採用選考・人事システムが徹底されること」、「若手研究者、女性研究者や海外の研究者の能力や業務が適切に評価されること」を期待する、としています。

なお、研究者等の業績の評価においても、やはりベースとなるのは被評価者である研究者が自ら実施する自己評価です。また、評価の重複を避けるために、「当該研究者等が関連する競争的資金制度における研究開発課題や国の実施する研究開発課題の評価の結果などを

適切に活用して効率的に実施する」ことも推奨されています。

## 2.2.6 研究開発機関等の評価

大学全般については「学校教育法」に基づく自己点検・評価や認証評価が、国立大学法人及び大学共同利用機関法人については「国立大学法人法」に基づく法人評価（教育研究の状況についての評価を含む）が、研究開発法人等については「独立行政法人通則法」に基づく法人評価が義務づけられています。

これらの評価の基本となるのは、機関自らが実施する評価です。文科省評価指針では、「研究開発活動に関する評価に関しては、機関の特性等に応じて、本指針を参考に、評価の目的や評価結果の活用の仕方、評価の項目・基準等を的確に設定し、実施することが期待される」としています。

なお、機関評価は、原理的には、機関を1つのプログラムとして見立てた「プログラム評価」の一種である、という見方も可能です。逆に言えば、機関の行う諸活動がミッションの達成に向けて適切にプログラム化されていないと、真に有効な機関評価は実施することができません。大学等にせよ国立研究開発法人等にせよ、機関評価において基本となるのはやはり機関の活動を構成する個々のプログラムを単位とした自己評価であり、プログラム評価を実質化することが組織経営上欠かせなくなってきました。

## 2.3 政策評価

行政機関<sup>6</sup>については、「行政機関が行う政策の評価に関する法律（政策評価法）」（平成13年6月29日法律第86号）に基づき、政策評価を行うことが義務付けられています。

政策評価制度は、「従来、わが国の行政においては、法律の制定や予算の獲得等に重点が置かれ、その効果やその後の社会経済情勢の変化に基づき政策を積極的に見直すといった評価機能は軽視されがちであった」（行政改革会議、平成9年12月3日）との認識の下、中央省庁等改革の1つの柱として始まりました。

この政策評価法に基づき、「政策評価に関する基本方針」（以下「基本方針」という。）が平成13年12月に閣議決定され、この方針に従って、各府省庁が3～5年を目処に「政策評価基本計画」を策定し、さらに年度毎に事後評価の実施に関する「政策評価実施計画」を策定して政策評価を行うこととなっています。

また、平成17年12月16日には、政策評価各府省連絡会議において、「政策評価の実施に関するガイドライン」が策定されました。これは、それまでの成果を踏まえて基本方針に基づき策定されたものであり、政策評価の円滑かつ効率的な実施のための標準的な指針を示したものです。

さらに、平成25年12月20日には、政策評価各府省連絡会議において、「目標管理型の政策評価の実施に関するガイドライン」が策定されました。その中で、「目的、目標（指標）、それらの達成手段、各手段がいかに目標等の実現に寄与するか等に係る事前の想定」を明確にするという趣旨の下、施策単位で「事前分析表」を作成することが示されました。なお、本ガイドラインでは、政策評価と行政事業レビューとの連携を確保するために、政策評価担

---

<sup>6</sup> 行政機関とは、政策評価法第二条に規定されている機関を指します。

当組織と行政事業レビューの取りまとめ部局との連携を行うことなどが記載されています。  
こうした基本方針やガイドライン等において、政策評価の具体的な内容や進め方等が規定されていますが、以下ではそのポイントについて解説します。

### 2.3.1 政策の階層

前述のように、「政策評価の実施に関するガイドライン」では、「政策評価の体系的かつ合理的で的確な実施を確保するため、政策体系をあらかじめ明らかにすること」が基本とされており、そのための区分として、「政策（狭義）」、「施策」、「事務事業」の3階層を導入しています。具体的には次のようなものです。

政策（狭義）：特定の行政課題に対応するための基本的な方針の実現を目的とする行政活動の大きなまとまり。

施策：上記の「基本的な方針」に基づく具体的な方針の実現を目的とする行政活動のまとまりであり、「政策（狭義）」を実現するための具体的な方策や対策ととらえられるもの。

事務事業：上記の「具体的な方策や対策」を具現化するための個々の行政手段としての事務及び事業であり、行政活動の基礎的な単位となるもの。

プログラムとは目的-手段の論理的な体系群であることは先に述べましたが、この観点からすれば、いずれの区分も大綱的指針で言う「プログラム」として扱うことができますし、「プログラム評価」の対象であるという見方もできます。うまくいけば、「プログラム」の概念を持ち込むことで、政策評価をアップグレードすることも可能になります。

一方、これらの区分は相対的なものであり、ある種の「理念」として示されたものにすぎません。ガイドラインにおいても、「現実の政策の態様は多様であることから、施策が複数の階層から成る場合や事務事業に相当するものが存在しない場合、一つの施策や事務事業が複数の政策体系に属する場合など、三つの区分に明確に分けることが困難なこともあり得る」としています。プログラムが有効に機能するためには論理的な体系としてまとまっていることが欠かせず、そのためには全体を構成する要素がモレなく、ダブリなくあげられている必要がありますが（MECE: Mutually Exclusive and Collectively Exhaustive）、上記で指摘されているような「ダブリ」などがある場合、評価に期待される改善効果も限定的になってしまう恐れがあります。プログラム評価を浸透させていく上でも、政策体系の論理化が今求められていると言えます。

### 2.3.2 3つの評価方式

政策評価法では、1) 事業評価方式、2) 実績評価方式、3) 総合評価方式という3つの評価が存在します。具体的には次のようなものです。

#### (1) 事業評価方式

事業評価方式については、基本方針において、「個々の事業や施策の実施を目的とする政策

を決定する前に、その採否、選択等に資する見地から、当該事業又は施策を対象として、あらかじめ期待される政策効果やそれらに要する費用等を推計・測定し、政策の目的が国民や社会のニーズ又は上位の目的に照らして妥当か、行政関与の在り方からみて行政が担う必要があるか、政策の実施により費用に見合った政策効果が得られるかなどの観点から評価するとともに、必要に応じ事後の時点で事前の時点に行った評価内容を踏まえ検証する方式」とされています（基本方針（別紙）「事業評価方式」）。

つまり、事業評価方式は、主に事務事業レベルの評価を扱うものであり、期待される効果や必要経費等を測る事前の時点での予測的評価と途上ないし事後の時点での検証とを、可能な限り広範に実施しようとするものです。

基本方針において、「対象とする政策が、どのような目的の下にどのような手段を用いるものかという対応関係を明らかにした上で行う」（基本方針 I-1-（2））とともに、「事後評価の実施に当たっては、行政目的と手段の関係を念頭に置きつつ、政策評価の結果を適切に反映するために合理的と認められる単位により行う」（同 I-5-イ）ものとされていますが、事業評価方式は、前述の実績評価方式と比べると、目的と手段との関係が明らかになりやすいという特徴をもっていると言えます。

なお、研究開発、公共事業及び政府開発援助の 3 分野における事業については、政策評価法の施行に先立ち、関係府省において評価について一定の取り組みが行われてきた経緯があることから、政策評価法では、事前評価における政策効果の把握の手法等が開発されている場合、これら 3 分野の個々の事業について事前評価の実施が義務付けられています。研究開発に関しては、10 億円以上の費用を要することが見込まれる「研究開発課題」がこれに該当します（評価法施行令第 3 条）。また、これら 3 分野の事業には長期間にわたって実施される大規模なものも多いことから、政策評価法第 7 条及び政策評価法施行令第 2 条各号の規定（政策の決定から 5 年間未着手又は 10 年間未了のものについての事後評価の実施）を踏まえ、多くの府省が事後評価を実施することとしています。

## (2) 実績評価方式

実績評価方式については、基本方針において、「政策を決定した後に、政策の不断の見直しや改善に資する見地から、あらかじめ政策効果に着目した達成すべき目標を設定し、これに対する実績を定期的・継続的に測定するとともに、目標期間が終了した時点で目標期間全体における取り組みや最終的な実績等を総括し、目標の達成度合いについて評価する方式」とされています（基本方針（別紙）「実績評価方式」）。

つまり、実績評価方式は政策の階層を問わず、あらかじめ設定した目標、とりわけ政策効果に着目した目標の達成度を、主として途上ないし事後の時点で定期的・継続的に測ろうとするものです。

政策効果とは、大綱的指針でいうところの「アウトカム」であり、政策評価法においては、「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及び社会経済に及ぼし、又は及ぼすことが見込まれる影響」と定義しています（政策評価法第 3 条第 1 項）。この政策効果に関して、政策評価法第 3 条第 2 項では、「できる限り定量的に把握すること」とされています。ただし、政策効果を定量的に把握することが「困難」であるかまたは「政策評価の客観的かつ厳格な実施の確保に結びつかない」場合には、定性的に把握する手法を用いることにはなりますが、その際にも、「できる限り、客観的な情報・データ

や事実を用いること」とされています（基本方針I-3-ア）。

### (3) 総合評価方式

総合評価方式については、基本方針において、「政策の決定から一定期間を経過した後を中心に、問題点の解決に資する多様な情報を提供することにより政策の見直しや改善に資する見地から、特定のテーマについて、当該テーマに係る政策効果の発現状況を様々な角度から掘り下げて分析し、政策に係る問題点を把握するとともにその原因を分析するなど総合的に評価する方式」とされています。

つまり、総合評価方式を用いた評価は、政策・施策を主な対象とし、途上時点を中心に特定テーマについての総合的な達成度を測ろうとするものであり、その政策手段とともに政策効果の発現状況を具体的に明らかにし、政策の問題点の把握とその原因の分析を実施するものであると言えます。

この他、政策評価法では総務省に他の省庁とは異なる立場からの評価機能を持たせており、総務省は複数の省にまたがる政策の評価、各府省庁の政策評価の点検活動を実施しています。点検の後、予算要求への評価結果の反映状況を取りまとめ、財務省がこれを受けて予算編成に活用します。

## 2.3.3 評価の内容及び方法

### (1) 評価の内容

政策評価法に基づく評価においては、上記のような対象に対し、次のような観点や一般的な基準を基本に総合的に評価を行うことになっています（「政策評価に関する標準的ガイドライン」）。

まず、評価の観点として重視されているのが、「必要性」、「効率性」、「有効性」です。「必要性」とは、「WHY」に関わるものであり、具体的には、「政策の目的が、国民や社会のニーズに照らして妥当か、上位の目的に照らして妥当か」、「行政関与の在り方から見て行政が担う必要があるか」といった問いに答えるものになります。これには、同じ目的で実施されている他の取り組みの中での位置づけを明確にするとともに、民間を含めた類似の取り組みとどのように差別化されているのかといった説明も含まれます。「効率性」とは、「HOW」に関わるものであり、評価においては、「投入された資源量に見合った効果が得られるか、又は実際に得られているか」「必要な効果がより少ない資源量で得られるものが他にないか」「同一の資源量でより大きな効果が得られるものが他にないか」といった問いに答える必要があります。「有効性」は、「WHAT」に関わるものであり、「政策の実施により、期待される効果が得られるか、又は実際に得られているか」といった目標設定の妥当性や実績を問うものです。

こうした3つの観点に加え、政策の性質によってはさらに「公平性」や、これらの観点からの評価を踏まえた「優先性」といった観点も標準的なガイドラインの中には示されており、各府省及び総務省は、実際に政策評価を実施するに当たっては、これらの一般基準を踏まえ、評価基準の具体化を図るものとする、とされています。

## (2) 評価の方法

また、評価手法に関しては、「各府省及び総務省は、評価に要するコスト等も勘案の上、評価の目的、評価対象の性質等に応じた適用可能で合理的な評価手法により政策評価を実施するものとする」とされています。その際、「まずは定量的な評価手法の開発を進めるよう努め、可能な限り具体的な指標・数値による定量的な評価手法を用いるよう努めること」や、「定量的な評価手法の適用が困難である場合又は客観性の確保に結び付かない場合などにおいては、定性的な評価手法を適用するものとし、その際、可能な限り、客観的な情報・データや事実に基づくものとし、評価において第三者等を活用するなどにより、評価の客観性の確保に留意する」必要があることは上に述べた通りです。

### 2.4 行政事業レビュー

行政の行うこうした政策を対象とした評価としては、その他にも「行政事業レビュー」があります。これは平成 22 年度から試行的に導入され、現在では「行政事業レビューの実施等について」（平成 25 年 4 月 5 日閣議決定）に基づき実施されています。次に掲げる取り組みを行うことで、「各府省庁が所掌する事業のより効果的かつ効率的な実施並びに国の行政に関する国民への説明責任及び透明性の確保を図り、もって国民に信頼される質の高い行政の実現を図る」ことを目指しています。

各府省庁自らが、事業に係る予算の執行状況等について、個別の事業ごとに整理した上で、毎会計年度終了後速やかに必要性、効率性及び有効性等の観点から検証して当該事業の見直しを行い、その結果を予算の概算要求及び執行に反映するとともに、それらの結果を公表すること。

また、こうした検証を行うにあたり、1) 事業に係る予算の執行状況等を分かりやすい形で公表すること、そして、2) 外部性が確保され、客観的かつ具体的で、厳格な検証となるよう、事業に係る予算の執行その他行政運営に関して識見を有する者の意見を聴くこと（この場合、効果的かつ効率的に意見聴取を行うとともに、一部の事業については、公開性を確保した上で行うこと）を求めています。また、各府省庁には、「補助金や出資等により造成された特定の基金を活用し、事業を実施している場合には、当該基金の執行状況等を分かりやすい形で毎年公表すること」も求められます。

なお、行政事業レビューの実施にあたっては、「統一的かつ効率的に実施する観点から、内閣官房行政改革推進本部事務局において、各府省庁に共通する手続の策定等を行い、これを推進すること」とされています。

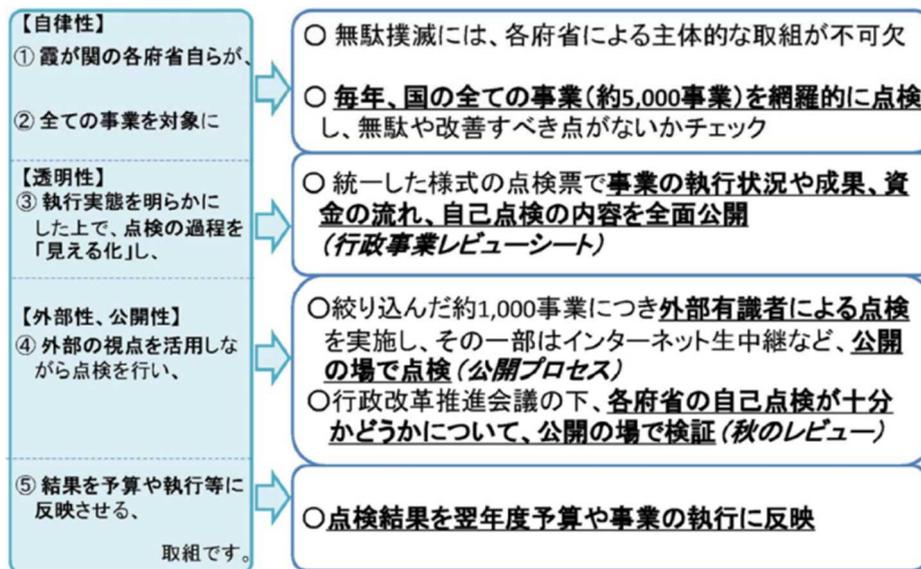


図 2-4 行政事業レビューの概要

出典:これでわかる！行政事業レビュー(平成 26 年度版)～PDCA で国の「仕事」を改善する～(平成 27 年 6 月 18 日内閣官房行政改革推進本部事務局)

## 2.5 法人評価

法人評価には、独立行政法人等を対象とするものと、国立大学法人および大学共同利用機関法人を対象とするものの 2 種類があります。以下では、それぞれについて解説します。

### 2.5.1 独立行政法人に対する法人評価

#### (1) 独立行政法人制度の枠組み

独立行政法人制度は、中央省庁等改革の一環として導入されたものであり、行政における企画立案部門と実施部門を分離し、企画立案部門の能力を向上させる一方で、実施部門に法人格を与え、運営裁量を与えることにより、政策実施のパフォーマンスを向上させることを目的として導入されたものです。公共性の高い事務・事業のうち、国が直接実施する必要はないが民間の主体に委ねると実施されないおそれのあるものを実施する法人がこれに該当します<sup>7</sup>。平成 13 年度からはじまったこの制度は、法的には独立行政法人の運営の基本その他の制度の基本となる共通の事項について定めた「独立行政法人通則法(以下、独法通則法)」(平成 11 年法律第 103 号)と、各独立行政法人の名称、目的、業務の範囲等に関する事項を定めた「個別法」により位置づけられています。

独立行政法人は国の政策を実現するための実施機関として各方面で成果をあげてきましたが、その運用の過程で様々な問題点も指摘され、第 1 次安倍内閣以降、国会等において制度改革をめぐる検討が積み重ねられてきました。平成 25 年 12 月には、「独立行政法人を国

<sup>7</sup> なお、独法通則法を準用する法人等として、国立大学法人や大学共同利用機関法人、日本私立学校振興・共済事業団(助成業務)等があります。

の施策の実施機関として明確に位置づけ、各々の法人が担う政策実施機能を最大限向上させるとともに、業務の質と効率を向上させるため、制度・組織面で抜本的な見直しを行う」ことを目的に、「独立行政法人改革等に関する基本的な方針」が閣議決定されました。この閣議決定に基づき、平成 26 年 6 月に成立したのが「独立行政法人通則法の一部を改正する法律」（平成 27 年 4 月施行）です<sup>8</sup>。

## (2) 独立行政法人制度の特徴

以下では、この基本方針および改正後の独法通則法の内容を中心に、制度改革の内容や現行の独立行政法人制度の特徴について解説します<sup>9</sup>。

### 1) 業務の特性を踏まえた法人の分類

法改正前は様々な分野で様々な態様の業務を行っている法人すべてを一律に規定していましたが、事務及び事業の特性に応じて、法人を「中期目標管理法法人」、「国立研究開発法人」、「行政執行法人」の 3 つに分類し、各分野に即した目標管理の仕組みを導入することとなりました。

まず、「中期目標管理法法人」とは、国民の需要に対応した多様で良質なサービスの提供を通じた公共の利益の増進を目的とするものであり、3～5 年の中期的な目標・計画に基づいて公共上の事務・事業を行う法人のことを指します（独法通則法第 29 条―第 35 条の 3）。文部科学省が所管する代表的な機関としては、日本学術振興会や大学改革支援・学位授与機構等があげられます。

2 番目の「国立研究開発法人」とは、科学技術に関する試験・研究・開発にかかわる業務を 5～7 年の中期的な目標・計画に基づいて行う法人です（独法通則法第 35 条の 4―第 35 条の 8）。「独立行政法人改革等に関する基本的な方針」では、研究開発型の法人についても他の独立行政法人と同様に「中期目標管理―評価」という枠組みが最適であるとしつつ、研究開発業務の特性を踏まえ、研究開発成果を最大化するという目的に適した仕組みが整備されることになりました（「独立行政法人改革等に関する基本的な方針」）。文部科学省が所管するこのタイプの法人としては、物質・材料研究機構、防災科学技術研究所、量子科学技術研究開発機構、科学技術振興機構、理化学研究所、宇宙航空研究開発機構、海洋研究開発機構、日本原子力研究開発機構の 8 機関になります。

「国立研究開発法人」のうち、物質・材料研究機構と理化学研究所、それから、経済産業省が所管する産業技術総合研究所の 3 機関は、「特定国立研究開発法人による研究開発等の促進に関する特別措置法」（平成 28 年法律第 43 号。以下、特定研発特措法）によって、「世界的な研究開発成果の創出を目指す法人」として「特定国立研究開発法人」に位置付けられています。なお、特定国立研究開発法人による研究開発等の促進の意義や基本的な方向、政府として講ずべき措置、体制の整備等に関しては、同法第 3 条第 1 項に基づき、「特定国立研究開発法人による研究開発等を促進するための基本的方針」（平成 28 年 6 月 28 日閣議決定）が定められています。

<sup>8</sup> 独立行政法人評価年報（平成 25 年度版）

<sup>9</sup> 「独立行政法人制度改革関連法案の骨子」（平成 26 年 4 月、内閣官房行政改革推進本部事務局）等を参照。

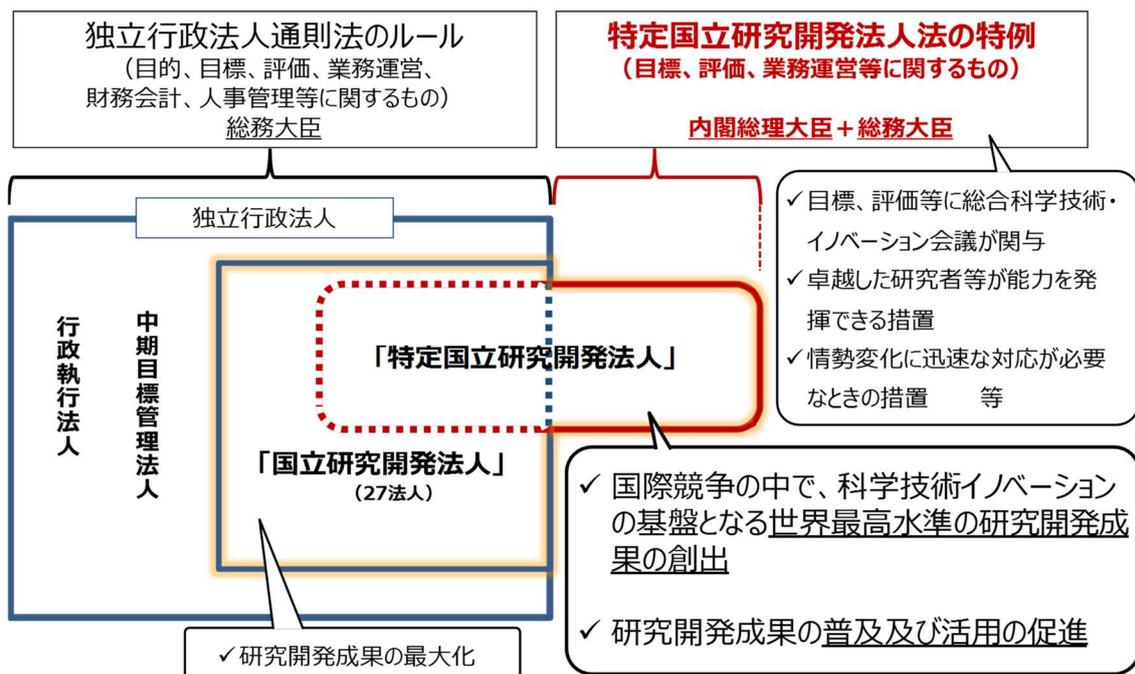


図 2-5 特定国立研究開発法人制度の概要

出典: 内閣府資料

最後の「行政執行法人」は、国の行政事務と密接に関連した事務・事業を、国の相当な関与のもとで単年度ごとの目標・計画に基づいて執行する法人であり、役職員には国家公務員の身分が付与されます。国立公文書館（内閣府所管）、統計センター（総務省所管）、造幣局（財務省所管）、国立印刷局（財務省所管）、農林水産消費安全技術センター（農林水産省所管）、製品評価技術基盤機構（経済産業省所管）、駐留軍等労働者労務管理機構（防衛省所管）がこれに該当します。

## 2) PDCA サイクルが機能する目標・評価の仕組みの構築

新たな制度では、主務大臣の下での政策の PDCA サイクルを強化するために、目標を指示する主務大臣が評価に関与していなかった制度を改め、主務大臣を評価主体とするなど目標・評価の一貫性・実効性を向上させることになりました。具体的には、主務大臣が与える目標の下で各法人の自主性・自律性が重視される一方、事後に主務大臣がその業務実績について評価を行い、業務・組織の見直しを図るといった仕組みです。

総務省では、こうした主務大臣による目標策定、業績評価が客観的かつ厳正に行われるよう、政府統一的な指針（「独立行政法人の目標の策定に関する指針」及び「独立行政法人の評価に関する指針」（平成 26 年 9 月 2 日総務大臣決定））を定めるとともに、政府唯一の第三者機関として独立行政法人評価制度委員会を設置し、主務大臣による目標策定、評価や業務・組織の見直しをチェックし、必要に応じて意見を述べることになっています。なお、総務省による 2 つの政府統一的な指針には、研究開発業務の持つ「長期性」「不確実性」「予見不可能性」「専門性」といった特性を踏まえ、総合科学技術・イノベーション会議によって作成された「研究開発成果の最大化に向けた国立研究開発法人の中長期目標の策定及び評価に関する指針」（平成 26 年 7 月 17 日総合科学技術・イノベーション会議決定）の内容が

反映されています。また、文部科学省では、こうして策定された 2 つの政府統一的な指針を受け、「文部科学省所管の独立行政法人の評価に関する基準」（平成 27 年 6 月 30 日文部科学大臣決定、平成 29 年 4 月 1 日一部改定）を策定し、それに基づいて目標・評価の仕組みを運用しています。

その他、国立研究開発法人については、上記のような研究開発業務の持つ特性を踏まえ、目標期間を最大 7 年とすることや、主務大臣による目標設定・評価を支援するための研究開発に関する審議会を設置することが改正後の独法通則法において定められています。さらに、特定国立研究開発法人については、「特定研発特措法」に基づき、目標設定や評価等に総合科学技術・イノベーション会議が関与することになっています。

### 3) その他の特徴

改正後の独法通則法では、監事の権限が不明確で、法人の違法行為に対して主務大臣から是正要求のみしか行えなかった制度を見直し、法人の内外から業務運営を改善し得るよう、法人内部のガバナンスを強化するほか、主務大臣による是正措置が整備されました。

また、中期目標管理法人・国立研究開発法人の役職員の報酬・給与等の基準は、国家公務員、民間企業、当該法人の業務の実績、職務の特性及び雇用形態その他の事情を考慮することになっています。

#### (3) 独立行政法人を対象とした法人評価の仕組み<sup>10</sup>

以上、独法通則法改正後の制度の特徴について解説しましたが、ここでは、中期目標管理法人と国立研究開発法人を対象を絞り、法人評価の仕組みを詳しくみていくことにします。

中期目標管理法人及び国立研究開発法人に関しては、共通して大きく次のような評価を行うことが独法通則法に定められています。なお、②及び③を実施する年度においては、①の評価もあわせて実施することになっています。

- ① 中（長）期目標期間中における事業年度ごとの業務実績評価（年度評価）
- ② 中（長）期目標期間の最後の事業年度の直前の事業年度時に行う中（長）期目標期間の終了時に見込まれる中（長）期目標期間全般の業務実績評価（見込評価）
- ③ 中（長）期目標期間の最後の事業年度に行う中（長）期目標期間全般の業務実績評価（期間実績評価）

この①～③の評価について、該当する各法人では、事業年度終了後 3 ヶ月以内に、業務の実績及びそれらについての自己評価の結果をまとめた報告書を主務大臣に提出、公表することになっています。また、主務大臣が①の年度評価を実施するにあたっては、中（長）期計画の実施状況の調査及び分析を行い、その結果を踏まえることとされています。

主務大臣によるこれら①～③の評価結果には総合的な評定が含まれ、それを各法人に通知、公表することになっていますが、その際、必要があると認められる場合には、法人に対して業務運営の改善等、必要な措置を講ずるよう命じることができます。特に、②の見込評価を実施した際には、主務大臣は、中（長）期目標期間の終了時までには当該法人の業務の継続や

<sup>10</sup> [http://www.soumu.go.jp/main\\_sosiki/hyouka/dokuritu\\_n/index.html](http://www.soumu.go.jp/main_sosiki/hyouka/dokuritu_n/index.html)

組織の存続の必要性など全般にわたる検討を行い、その結果に基づいて、業務の廃止や移管、組織の廃止等の措置を講ずることになっています。これらの結果は当該法人の次期中（長）期目標の設定にも反映されます。なお、この見込評価については、総務省に設置の独立行政法人評価制度委員会（以下、委員会）に対しても結果の通知を行い、それを受けた委員会は必要に応じて主務大臣に対して意見を述べることとされています。

上記の仕組みは中期目標管理法及び国立研究開発法人に共通するものですが、国立研究開発法人については、研究開発の特性を踏まえ、研究開発の事務及び事業に関する事項について、あらかじめ研究開発に関する審議会の意見を聴くこととされています。主務大臣が②の見込評価時に行う一般的な検討においても同様です。その他、中長期目標の期間よりも当該法人の長の任期が短い場合、任期期間中の事業年度全般における業務の実績について、主務大臣の評価を受けることになっています。

さらに、国立研究開発法人のうち特定国立研究開発法人については、国家戦略の実現という観点から、特定研発特措法に基づき、主務大臣は中長期目標の策定・変更、見込評価、中長期目標期間の終了時の見直し等に際して総合科学技術・イノベーション会議に意見を聴かなければならないとされています。また、大綱的指針では、「主務大臣が行う特定国立研究開発法人の中長期目標期間終了時に見込まれる業績評価及び終了時の検討結果を参照し、科学技術・イノベーション政策に適時適切に反映させるほか、必要に応じて研究開発に係る評価等に関する指針類の見直しを行う」ことが述べられています。

なお、「独立行政法人の評価に関する指針」を基に定められた「文部科学省所管の独立行政法人の評価に関する基準」においては、以上のような独法通則法等の法律や閣議決定された政府の方針等による評価に加え、法人の業務等に関する国会審議、会計検査、予算執行調査等の指摘事項への取り組み事項についても評価を実施することが明記されています。また、主務大臣の下での PDCA サイクルを十分に機能させるという制度改革の趣旨を踏まえ、各法人による業務実績に係る自己評価に加え、関連する国の政策評価、行政事業レビュー、そして行政評価・監視の結果を活用して評価を行うこととされています。

#### (4) 独立行政法人を対象とした法人評価の内容

独立行政法人の法人評価は、前述のように、各法人による自己評価を基本としていますが、その内容についてみると、原則「独立行政法人の目標の策定に関する指針」に基づき定められた中（長）目標の項目を評価単位として行う「項目別評定」と、これを基礎とした上で法人全体を評価する「総合評定」から成り立っています。

目標に定めるべき事項とは、次のようなものです。

表 2-4 独立行政法人の法人評価における目標に定めるべき事項

中期目標管理法 (独法通則法第 29 条第 2 項)	国立研究開発法人 (独法通則法第 35 条の 4 第 2 項)
<ul style="list-style-type: none"> <li>● 国民に対して提供するサービスその他の業務の質の向上に関する事項</li> <li>● 業務運営の効率化に関する事項</li> <li>● 財務内容の改善に関する事項</li> <li>● その他業務運営に関する重要事項</li> </ul>	<ul style="list-style-type: none"> <li>● 研究開発の成果の最大化その他の業務の質の向上に関する事項</li> <li>● 業務運営の効率化に関する事項</li> <li>● 財務内容の改善に関する事項</li> <li>● その他業務運営に関する重要事項</li> </ul>

出典：未来工学研究所作成

これらの各項目に対する評価の具体的な指針や基準等は、「独立行政法人の目標の設定に関する指針」や「文部科学省所管の独立行政法人の評価に関する基準」に定められています。

項目別評定にせよ総合評定にせよ、目標及び計画で掲げる指標を基準とする絶対評価によって行うものとされており、Bを標準としてS、A、B、C、Dの5段階（これを「評語」と呼んでいます）で表現することになっています。なお、研究開発に係る事務及び事業については、研究開発業務の特性等を踏まえ、目標設定時に適切な評価軸を設定することとされています。

## 2.5.2 国立大学法人等に対する法人評価

### (1) 法人制度の概要

#### 1) 法人化の経緯と法的位置づけ

国立大学及び大学共同利用機関は平成16年4月に法人化され、それぞれ国立大学法人及び大学共同利用機関法人と呼ばれるようになりました（この両者はあわせて「国立大学法人等」と呼ばれます）。国立大学法人等は、平成15年に制定された国立大学法人法（平成15年法律第112号）によってその組織や運営のあり方が定められており、また、一部前述の独法通則法の規定を準用しています。

法人化の背景には、「大学としての特性を踏まえて様々な特例措置を講じてはいるものの、基本的には行政組織の一部として位置付けられて」きたために、「国の予算制度や国家公務員法制の下で、教育研究の柔軟な展開に制約」があったという問題意識があります<sup>11</sup>。こうした問題を克服し、国立大学が国民や社会の期待にこたえてその役割を一層しっかりと果たしていくためには、各大学の運営上の裁量を大幅に拡大し、より自律的な環境の下で教育研究の活性化と質の向上を図ることが必要であるという考えに基づき、それぞれが法人格を持った独立した法人とすることになりました。具体的には、以下のような基本的な考え方のもと、抜本的な改革を進めることになりました。

- 「大学ごとに法人化」し、自律的な運営を確保すること
- 「民間的発想」のマネジメント（運営管理）手法を導入すること
- 「学外者の参画」による運営システムを制度化すること
- 国家公務員法体系にとらわれない非公務員型の弾力的な人事システムへ移行すること
- 評価の実施による事後チェック方式に移行すること

#### 2) 中期目標及び中期計画の策定プロセス

こうした改革を実現するために、各国立大学法人等では6年間の中期目標及び中期計画を定め、それに基づいて組織経営を行うことになりました。平成30年3月現在、第3期中期目標期間にあたります。

まず、中期目標については、各法人が作成した原案をもとに国立大学法人評価委員会が審議を行い、その意見を踏まえて、文部科学大臣が各法人に中期目標を提示し、公表する、と

---

<sup>11</sup> 平成15年度文部科学白書

いう流れになっています。これを変更する場合も同様です。国立大学法人評価委員会は、国立大学法人法第 9 条に基づいて文部科学省に設置されているものであり、政令の定めるところにより、文部科学大臣が任命した 20 人以内の学識経験者・有識者で構成されています。なお、平成 28 年 5 月に成立した「国立大学法人法の一部を改正する法律」を受け、新たに指定国立大学法人制度がはじまりましたが、平成 29 年 4 月以降、文部科学大臣が指定国立大学法人における中期目標の設定及び変更を行うにあたっては「世界最高水準の教育研究活動を行う外国の大学の業務運営の状況を踏まえなければならない」ことになりました（国立大学法人法第 34 条の 6）。

さて、この中期目標では、以下の 5 項目について定めることになっています（国立大学法人法第 30 条第 2 項）。

- 教育研究の質の向上に関する事項
- 業務運営の改善及び効率化に関する事項
- 財務内容の改善に関する事項
- 教育及び研究並びに組織及び運営の状況について自ら行う点検及び評価並びに当該状況に係る情報の提供に関する事項
- その他業務運営に関する重要事項

中期計画は、こうした中期目標を達成するために各法人が作成するものであり、これも国立大学法人評価委員会が審議を行い、その意見を踏まえて文部科学大臣が認可を行います。これを変更する場合も同様の手続をとります。中期計画においては、上記の中期目標で掲げた事項を達成するために、以下のような事項を定めることになっています（国立大学法人法第 31 条第 2 項）。

- 教育研究の質の向上に関する目標を達成するためとるべき措置
- 業務運営の改善及び効率化に関する目標を達成するためとるべき措置
- 予算（人件費の見積りを含む。）、収支計画及び資金計画
- 短期借入金の限度額
- 余剰金の使途
- 重要な財産を譲渡し、又は担保に供しようとするときは、その計画
- その他文部科学省令で定める業務運営に関する事項

なお、文部科学大臣は、中期目標に掲げる事項の「適正かつ確実な実施上不適当となったと認める」場合、「その中期計画を変更すべきことを命ずることができる」とされています（国立大学法人法第 31 条 4 項）。

以上、中期目標及び中期計画の策定プロセスについてみてきましたが、国立大学法人評価委員会が行うこうした審議は実質的には機関の事前評価にあたるという見方もできます。

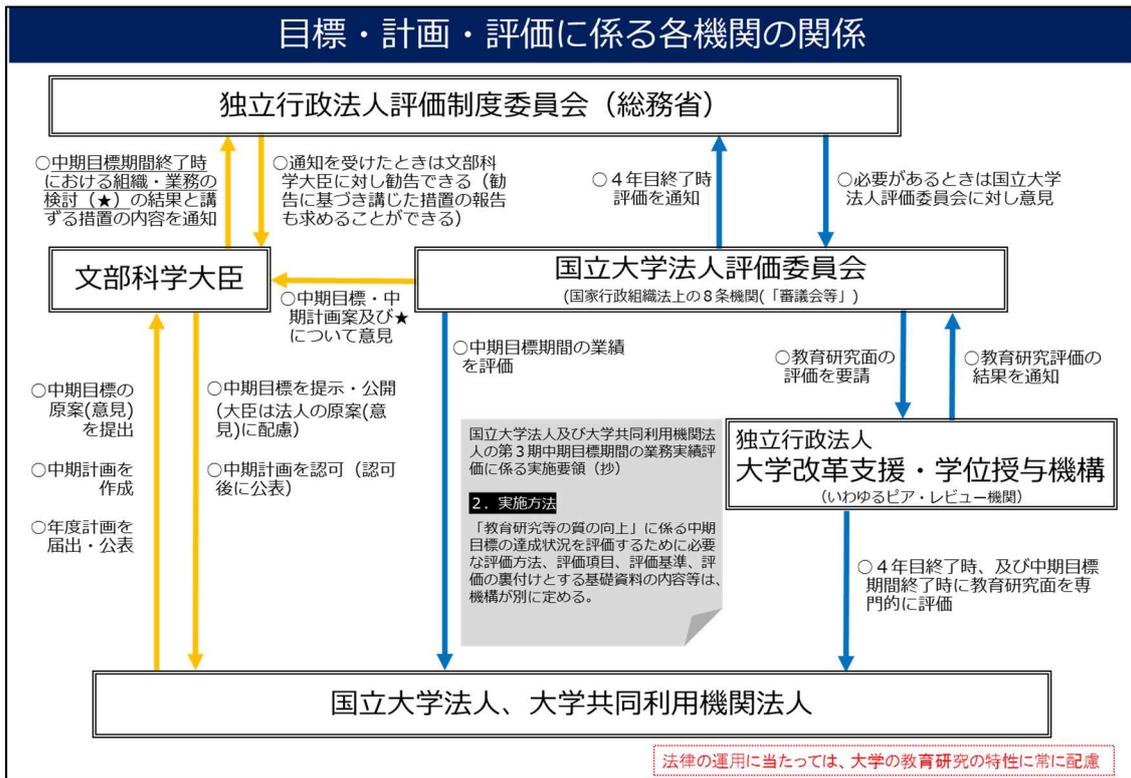


図 2-6 国立大学法人等の目標・計画・評価に係る各機関の関係

出典：文部科学省資料

## (2) 国立大学法人等を対象とした法人評価の仕組み

国立大学法人等の法人評価は独法通則法及び国立大学法人法に基づいて実施されますが、各法人の意向に最大限配慮すること、そして、国立大学法人評価委員会及び第三者評価機関である大学改革支援・学位授与機構が関わることにその特徴があります。

具体的には、独立行政法人の法人評価と同様、次のような評価を行います。なお、②及び③を実施する年度においては、①の評価もあわせて実施することになっています。

- ① 中期目標期間中における事業年度ごとの業務実績評価（年度評価）
- ② 中期目標期間の最後の事業年度の前々事業年度時に行う中期目標期間の終了時に見込まれる中期目標期間全般の業務実績評価（見込評価）
- ③ 中期目標期間の最後の事業年度に行う中期目標期間全般の業務実績評価（期間実績評価）

この①～③の評価において、各法人では、事業年度終了後 3 ヶ月以内に、業務の実績及びそれらについての自己評価の結果をまとめた報告書を国立大学法人評価委員会に提出、公表することになっています。また、国立大学法人評価委員会が①～③の評価を実施するにあたっては、中期計画の実施状況の調査及び分析を行い、その結果を踏まえることとされています。さらに、②及び③の評価にあたっては、国立大学法人法第 31 条の 3 第 1 項に基づき、大学改革支援・学位授与機構に対して「大学等の教育研究活動等の状況」の評価を実施するよう要請し、その評価結果を尊重することになっています。なお、大学改革支援・学位授与機構による評価の結果は、当該法人等にも提供され、公表されます。

国立大学法人評価委員会によるこれら①～③の評価結果には総合的な評定が含まれ、それを各法人に通知、公表することになっていますが、その際、必要があると認められる場合には、法人に対して業務運営の改善その他の勧告をすることができます。特に、国立大学法人評価委員会が②の見込評価を実施した際には、総務省に設置の独立行政法人評価制度委員会に対しても通知を行います。評価制度委員会は、その結果について、必要があると認める場合には国立大学法人評価委員会に対して意見を述べるすることができます。また、文部科学大臣は、この見込評価が実施された時には、中期目標期間の終了時までには当該法人の業務を継続させる必要性や組織のあり方など全般にわたる検討を行い、その結果に基づいて、また、国立大学法人評価委員会の意見を踏まえた上で、当該法人に関し所要の措置を講ずることになっています。なお、文部科学大臣は、この見込評価に基づく検討及びそれに基づき講じる措置の内容について、独立行政法人評価制度委員会に対して通知を行います。通知を受けた同委員会は、中期目標の期間の終了時までには、当該法人の主要な事務及び事業の改廃に関して、文部科学大臣に勧告することができます。

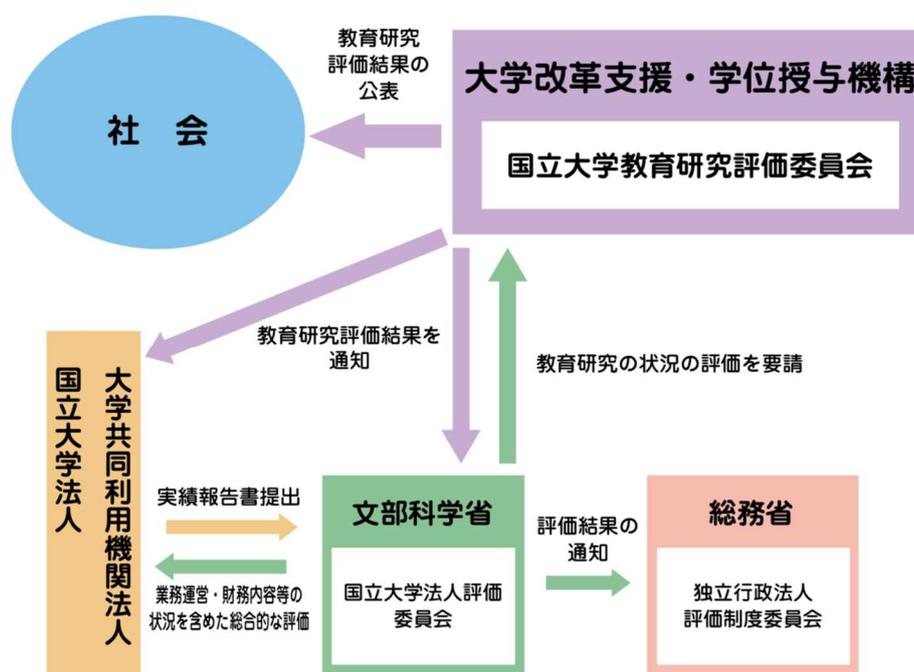


図 2-7 中期目標期間評価の仕組み

出典：大学改革支援・学位授与機構概要 2017

### (3) 国立大学法人等を対象とした法人評価の内容

以上みてきたように、法的に要求される国立大学法人等の法人評価では、各法人による自己評価が最初に求められます。これを用いた法人評価の実施方法等が「国立大学法人及び大学共同利用機関法人の第 3 期中期目標期間における各年度終了時の評価に係る実施要領」（平成 27 年 5 月 27 日国立大学法人評価委員会決定。平成 29 年 3 月 2 日一部改正）と、「国立大学法人及び大学共同利用機関法人の第 3 期中期目標期間の業務実績評価に係る実施要領」（平成 27 年 5 月 27 日国立大学法人評価委員会決定）にまとめられています。

「自己点検・評価」の詳細については続く 2.6 でとりあげることとし、以下では、2 つ

の実施要領に基づき、法人評価の内容をまとめます。なお、独立行政法人の法人評価では大きく「項目別評定」と「総合評定」の2項目から構成されていましたが、国立大学法人等の場合もこれと同様に、「項目別評価」と「全体評価」の2つから成り立っています。

## 1) 年度評価の内容

まず、年度評価における項目別評価の対象となるのは、前述の中期目標に定めるべき5項目のうち、1) 業務運営の改善及び効率化、2) 財務内容の改善、3) 自己点検・評価及び情報提供、4) その他業務運営（施設設備の整備・活用、安全管理及び法令遵守等）の4項目に関する取組状況であり、これらをまとめて「業務運営・財務内容等の状況」と呼んでいます。これらについて、年度計画の記載事項ごとに事業の実施状況を自己点検・評価し、実績報告書に次のような4段階で進捗状況を記載することになっています。なお、「産業競争力強化法」（平成25年法律第98号）が成立したことに伴い、「産業競争力強化法の規定による出資等に関する目標の状況」についても、同様の方法により記載することになりました。

表 2-5 自己点検・評価における項目別評価の評定区分

進捗状況	記号
年度計画を上回って実施している	(Ⅳ)
年度計画を十分に実施している	(Ⅲ)
年度計画を十分には実施していない	(Ⅱ)
年度計画を実施していない	(Ⅰ)

中期目標に定めるべき項目のうち残り1つに関わる「教育研究等の質の向上の状況」については、これらの区分によらず、実績報告書の「全体的な状況」欄に総括して記載します。

こうした各法人による実績報告書を受け、国立大学法人評価委員会では、書面調査及びヒアリングを通じて、各法人において自己点検・評価が着実に進められているかどうかを確認するとともに、各法人の年度計画の実施状況等に基づいて、中期計画の各項目の進捗状況を確認（項目別評価）し、その結果等を踏まえて、各法人の特性に配慮しながら、中期計画の達成に向けた進捗状況の総合的な評価（全体評価）を行います。

評価委員会による項目別評価では、6段階で絶対評価を行い、以下の判断基準を目安に各法人の諸事情に勘案して総合的に判断することになっています（表 2-6）。「産業競争力強化法の規定による出資等に関する目標の状況」についても、同様の方法により記載します。

「教育研究等の質の向上の状況」については、各法人による実績報告書の「全体的な状況」欄の記載などを踏まえて、注目すべき点にコメントをつけます。

評価委員会による全体評価については、項目別評価結果及び実績報告書の「全体的な状況」欄の総括的な記述等を踏まえ、各法人の特性に配慮しつつ、各法人が中期目標の前文に掲げている「基本的な目標」に対する具体的な取組状況及び中期計画の進捗状況を記述式で評価します。なお、「戦略性が高く意欲的な目標・計画」については、各法人の質的向上を促すという観点から、達成状況に加えプロセスや内容を評価するなど積極的な取組として適切に評価することになっています。

こうして実施された評価結果の案は、その透明性・正確性を確保するために、評価結果を確定する前に各法人に示されます。そして、各法人はそれに対し、意見申し立てを行うことができる、という制度になっています。

表 2-6 評価委員会における項目別評価の評定区分及び判断基準（目安）

評定	判断基準(目安)
中期計画の達成に向けて特筆すべき進捗状況にある	国立大学法人評価委員会が特に認める場合
中期計画の達成に向けて順調に進んでおり一定の注目事項がある	すべてⅣ又はⅢかつ注目すべき点が一定程度ある場合
中期計画の達成に向けて順調に進んでいる	すべてⅣ又はⅢ
中期計画の達成に向けておおむね順調に進んでいる	Ⅳ又はⅢの割合が9割以上
中期計画の達成のためには遅れている	Ⅳ又はⅢの割合が9割未満
中期計画の達成のためには重大な改善事項がある	国立大学法人評価委員会が特に認める場合

## 2) 見込評価及び期間実績評価の内容

中期目標期間における業務実績の評価、すなわち、見込評価及び期間実績評価においても項目別評価と全体評価を行うことに変わりはありませんが、評価の内容や方法、手続が年度評価の場合とはいくつかの点で異なってきます。

まず、中期目標期間における業務実績の評価では、項目別評価においても「教育研究等の質の向上」に関する中期目標の達成状況を確認します。その際、各法人が後述の自己点検・評価や認証評価のために整えた根拠資料やデータ等を活用し、評価作業の合理化を図ることとされていますが、実質的な評価を担うのは第三者機関である大学改革支援・学位授与機構です。国立大学法人評価委員会が評定を行う際には、同機構の評価結果を尊重することになっており、評価方法、評価項目、評価基準、評価の裏づけとする基礎資料の内容等も、同機構が別に定めることになっています。

同機構が行う評価のうち、基本となるのが学部・研究科等の単位で行う「教育研究の質の向上の状況を含む水準の評価」です。同機構では、各法人の自己点検・評価の内容等を検証した上で、こうした教育研究の質の向上の状況を含む水準の評価結果を勘案し、「教育に関する目標」、「研究に関する目標」、「社会との連携や社会貢献及び地域を志向した教育・研究に関する目標」（大学共同利用機関法人については「社会との連携及び社会貢献に関する目標」）、「その他の目標」、「共同利用・共同研究に関する目標」（大学共同利用機関法人のみ）といった項目ごとに、中期目標の達成状況を5段階で評定します（表 2-7）。その際、優れた点や改善すべき点を、各法人の自主的な教育研究等の改善に資する観点から分かりやすく指摘することになっています。これらの評価結果の決定に先立っては、評価の透明性・正確性を確保するため、各法人に評価結果案に対する意見申立ての機会が付与されます。

表 2-7 大学改革支援・学位授与機構における「教育研究等の質の向上」評価の評定区分

評定
中期目標の達成状況が非常に優れている
中期目標の達成状況が良好である
中期目標の達成状況がおおむね良好である
中期目標の達成状況が不十分である
中期目標の達成のためには重大な改善事項がある

なお、同機構が実施する上記評価のうち、「その他の目標」に関わる内容には、「附属病院に関する目標」、「附属学校に関する目標」、「産業競争力強化法の規定による出資等に関する

目標」は含まれません。これらについては、各法人が行う自己点検・評価に基づき、その妥当性を含めて、国立大学法人評価委員会が総合的に検証することになっています。その際、「附属病院に関する目標」、「附属学校に関する目標」については、その特性に配慮し、同機構が行う学部・研究科等の教育研究の質の向上の状況を含む水準の評価結果を参考に検証を行います。これらの評定も各法人の中期目標の達成状況に対する絶対評価で行いますが、詳細については割愛します。

「業務運営・財務内容等の状況」については、基本的に年度評価と同様、1) 業務運営の改善及び効率化、2) 財務内容の改善、3) 自己点検・評価及び情報提供、4) その他業務運営（施設設備の整備・活用、安全管理及び法令遵守等）の4項目に関する取組状況を中期計画の記載事項ごとにみていくこととなりますが、評価委員会が評定を行うにあたっては、各法人が共通に取り組む必要があり、かつ評価委員会が特に確認する必要がある項目を勘案することになっています。「共通の観点」と称されるこれらの項目は、ガバナンス改革、財務内容の改善、法令遵守及び研究の健全化の3点からなり、具体的には次のようなものです。

表 2-8 「共通の観点」の概要

	観点	概要
1	ガバナンス改革	権限と責任が一致した意思決定システムの確立や法人運営組織の役割分担の明確化、監事の役割の強化等のガバナンス改革の取組が、人材育成・イノベーションの拠点としての法人の機能強化にどのようにつながっているかという観点から評価。また、外部有識者の意見が法人運営に適切に反映され法人運営の活性化につながっているかという観点からも併せて評価。 (確認事項例) ・ 戦略的・効果的な法人運営・資源配分の仕組みとその効果 ・ 内部監査や監事監査結果の法人運営への反映状況
2	財務内容の改善	外部資金の一層の獲得や財源の多様化等による自己収入の増加を図るとともに、財務データを最大限活用し、事実関係に裏打ちされた財務分析を実施することで、資源配分の重点化や経費の節減などその分析結果を運営の改善に活用しているかどうかという観点から評価。 (確認事項例) ・ 既定収入の見直しや新たな収入源の確保に向けた取組状況 ・ 財務情報に基づく財務分析結果の活用状況
3	法令遵守及び研究の健全化	法人が社会的使命を果たしつつ、その活動を行っていくため、法令遵守(コンプライアンス)及び危機管理体制が具体的にどのように機能しているかという観点から評価。法人が研究の健全化のために、研究費の不正使用や研究活動における不正行為を防止するための体制が具体的にどのように機能しているかという観点から評価。 (確認事項例) ・ 法令遵守に関する体制及び規程等の整備・運用状況 ・ 災害、事件・事故等の危機管理に関する体制及び規程等の整備・運用状況 ・ 研究者及び学生に対する研究倫理教育の実施状況

出典:国立大学法人及び大学共同利用機関法人の第3期中期目標期間の業務実績評価に係る実施要領(別添1)をもとに未来工学研究所作成

なお、中期目標期間における業務実績の評価においても、年度評価と同様、各法人による自己点検・評価、国立大学法人評価委員会による検証及び評定といった流れで、表 2-9 及

び表 2-10 の評定区分等に基づいて、項目ごとの絶対評価を行います。なお、こうした評定区分等については適宜見直しがされており、実際に評価を行うに当たっては最新の情報を入力するよう心がけていただくことをおすすめします。

表 2-9 自己点検・評価における「業務運営・財務内容等の状況」評価の評定区分

進捗状況	記号
中期計画を上回って実施している	(Ⅳ)
中期計画を十分に実施している	(Ⅲ)
中期計画を十分には実施していない	(Ⅱ)
中期計画を実施していない	(Ⅰ)

表 2-10 評価委員会における「業務運営・財務内容等の状況」評価の評定区分及び判断基準（目安）

評定	判断基準(目安)
中期目標の達成状況が非常に優れている	国立大学法人評価委員会が特に認める場合
中期目標の達成状況が良好である	すべてⅣ又はⅢ
中期目標の達成状況がおおむね良好である	Ⅳ又はⅢの割合が9割以上
中期目標の達成状況が不十分である	Ⅳ又はⅢの割合が9割未満
中期目標の達成のためには重大な改善事項がある	国立大学法人評価委員会が特に認める場合

評価委員会では、以上のような中期目標各項目に対する項目別評価の結果等を踏まえつつ、各法人の特性にも配慮し、中期目標期間の業務実績の全体を記述式によって評価します。また、各法人が中期目標の前文に掲げている「基本的な目標」や、「戦略性が高く意欲的な目標・計画」に対する具体的な取組状況についても記述することになっています。

こうして実施された評価結果の案は、その透明性・正確性を確保するために、評価結果を確定する前に各法人に示されます。そして、各法人はそれに対し、意見申し立てを行うことができる、という制度になっています<sup>12</sup>。

## 2.6 自己点検・評価／認証評価

### 2.6.1 自己点検・評価及び認証評価の導入経緯と法人評価との関係

自己点検・評価及び認証評価は、学校教育法に基づいて実施されるものであり、それぞれ次のように規定されています。

大学は、その教育研究水準の向上に資するため、文部科学大臣の定めるところにより、当該大学の教育及び研究、組織及び運営並びに施設及び設備（次項において「教育研究等」という。）の状況について自ら点検及び評価を行い、その結果を公表するものとする。（学校教育法第 109 条第 1 項）

<sup>12</sup> 大学改革支援・学位授与機構が専門的な観点から行う教育研究の状況の評価に関しては、同機構において、評価結果の決定に先立ち、各法人に意見申立ての機会を付与した上で評価結果を決定していることから、意見申立ての対象にはなりません。

大学は、前項の措置に加え、当該大学の教育研究等の総合的な状況について、政令で定める期間ごとに、文部科学大臣の認証を受けた者(以下「認証評価機関」という。)による評価(以下「認証評価」という。)を受けるものとする。(学校教育法第 109 条第 2 項)

大学における自己点検・評価は、平成 3 年の大学設置基準の大綱化に伴い実施が努力義務化され、平成 11 年の大学設置基準の改正によって実施と結果の公表が義務化、結果の学者による検証が努力義務化されました。その後、平成 14 年の学校教育法の改正により、自己点検・評価の実施と結果の公表に関する規定が法律上明示され、平成 16 年度から施行されています。また、学校教育法の改正により、すべての大学が平成 16 年 4 月から文部科学大臣の認証を受けた認証評価機関による評価を受けることになりました。平成 16 年度はちょうど国立大学が法人化されたタイミングでもあり、ここに、自己点検・評価、認証評価、法人評価といった現行の大学評価の仕組みが整うこととなります。

## 2.6.2 自己点検・評価

自己点検・評価の目的は、「大学が、教育研究水準の向上や活性化に努めるとともに、その社会的責任を果たしていくため、その理念・目標に照らして自らの教育研究活動等の状況について自己点検し、現状を正確に把握・認識した上で、その結果を踏まえ、優れている点や改善を要する点など自己評価を行うこと」(中央教育審議会大学分科会資料)です。各大学の理念・目標は、例えば国立大学法人の場合、6 年ごとに改定される「中期目標」の中で基本的な目標や項目別の目標として表現されていますが、自己点検・評価は、こうした大学としての目指す姿に照らした進捗や実績などの現状把握に基礎におきながら、社会に対して説明責任を果たすとともに、自らの活動を改善するために行うものであることが分かります。

自己点検・評価には決まった項目はなく、法的に定められた周期などもあるわけではありませんが、大学審議会の答申「大学教育の改善について」(平成 3 年 2 月 8 日)では、「大学の自己点検・評価項目(例)」も示されています。研究開発評価に関わりの深い主な部分を抜き出すと、次のようなものです。

表 2-11 大学の自己点検・評価項目(例)(抜粋)

◇教育理念・目標等
<ul style="list-style-type: none"> <li>○ 大学(学部)の教育理念・目標の設定</li> <li>○ 教育理念・目標の点検・見直し</li> <li>○ 大学(学部)の将来構想</li> <li>○ 教育研究の活性化・充実のためのこれまでの取り組み etc.</li> </ul>
◇教育活動
(学生の受け入れ)／(学生生活への配慮)／(カリキュラムの編成)／(教育指導の在り方)／(教授方法の工夫・研究)／(成績評価, 単位認定)／(卒業生の進路状況)
◇研究活動
<ul style="list-style-type: none"> <li>○ 構成員による研究成果の発表状況</li> <li>○ 研究誌の発行状況と編集方針</li> <li>○ 共同研究の実施状況</li> <li>○ 研究費の財源(学外からの資金の導入状況, 科学研究費補助金の採択状況等)</li> <li>○ 研究費の配分方法</li> <li>○ 学会活動への参加状況 etc.</li> </ul>

◇教員組織	<ul style="list-style-type: none"> <li>○ 採用, 昇進の手順・基準</li> <li>○ 教員の兼職の方針と状況</li> <li>○ 教員人事についての長期計画 etc.</li> </ul>
◇施設設備	<ul style="list-style-type: none"> <li>○ 施設設備の整備・運用状況</li> <li>○ 学術情報システムの整備・活用状況 etc.</li> </ul>
◇国際交流	<ul style="list-style-type: none"> <li>○ 留学生の受け入れ状況(受け入れ数, 奨学金, 宿舍等), 指導体制</li> <li>○ 在学生の海外留学・研修の方針と状況</li> <li>○ 教員の在外研究の方針と状況</li> <li>○ 海外からの研究者の招致状況</li> <li>○ 海外の大学との交流協定の締結状況と活用状況 etc.</li> </ul>
◇社会との連携	<ul style="list-style-type: none"> <li>○ 教員の学外活動状況</li> <li>○ 学外の意見を教育研究に反映させるしくみ etc.</li> </ul>
◇管理運営, 財政	<ul style="list-style-type: none"> <li>○ 教育研究に関する意志決定の方法・体制</li> <li>○ 事務組織</li> <li>○ 予算の編成と執行の方針と状況</li> </ul>
◇自己評価体制	<ul style="list-style-type: none"> <li>○ 学外資金の導入状況 etc.</li> <li>○ 自己評価を行うための学内組織</li> <li>○ 教育研究活動等の公表</li> <li>○ 評価をフィードバックするためのしくみ etc.</li> </ul>

出典: 大学審議会の答申「大学教育の改善について」(平成 3 年 2 月 8 日)をもとに未来工学研究所作成

また、上記答申が出されてから様々な評価制度が制度され、法人評価や認証評価にも自己点検・評価作業が大学に求められるようになりました。法律の上では、法人評価や認証評価において求められている自己評価が、学校教育法が定める自己点検・評価に当たるのかは明確になっておりません。そのため、多くの大学では法人評価や認証評価で求められる内容や周期と調和するよう評価を実施しています。

### 2.6.3 認証評価

認証評価制度は、国公立すべての大学を対象とするものであり、当該大学の理念・目標に照らし、教育研究、組織運営及び施設設備の総合的な状況について、7 年以内ごとに、文部科学大臣が認証する評価機関（認証評価機関）による評価を受けることを各大学に義務付けるものです<sup>13</sup>。その大きな目的は、「大学の質の保証」にあります。具体的には、各大学の教育研究水準の向上に資するため、第三者機関である「認証評価機関が評価結果を公表することにより、大学が社会による評価を受ける」こと、そして、「評価結果を踏まえて大学が自ら改善することを促す」ことをねらいとしています（中央教育審議会大学分科会資料）。

大学の認証評価機関としては表 2-12 に示した 3 機関があります。評価はこれらの認証評価機関が独自に定めている大学評価基準に従って行われます。各認証評価機関は独自の特徴を

<sup>13</sup> 専門職大学院については、5 年に 1 回の周期になります。

持っており、各大学はこれら複数の機関の中から選択して評価を受けることが可能です。そのため、多元的な評価制度であると言われています。

表 2-12 認証評価機関における評価基準（比較）

	大学基準協会 <sup>14</sup>	大学改革支援・学位授与機構 <sup>15</sup>	日本高等教育評価機構 <sup>16</sup>
法人種別	公益財団法人	独立行政法人	公益財団法人
認証年月日	平成 16 年 8 月 31 日	平成 17 年 1 月 14 日	平成 17 年 7 月 12 日
評価名称	大学評価	大学機関別認証評価	大学機関別認証評価
評価基準	1 理念・目的 2 教育研究組織 3 教員・教員組織 4 教育内容・方法・成果 5 学生の受け入れ 6 学生支援 7 教育研究等環境 8 社会連携・社会貢献 9 管理運営・財務 10 内部質保証	基準 1 大学の目的 基準 2 教育研究組織 基準 3 教員及び教育支援者 基準 4 学生の受入 基準 5 教育内容及び方法 基準 6 学習成果 基準 7 施設・設備及び学生支援 基準 8 教育の内部質保証システム 基準 9 財務基盤及び管理運営 基準 10 教育情報等の公表	基準 1 使命・目的等 基準 2 学修と教授 基準 3 経営・管理と財務 基準 4 自己点検・評価

出典：各種資料をもとに未来工学研究所作成

なお、認証評価のプロセスについて、大学改革支援・学位授与機構を例にとると、図 2-8 のようになります。

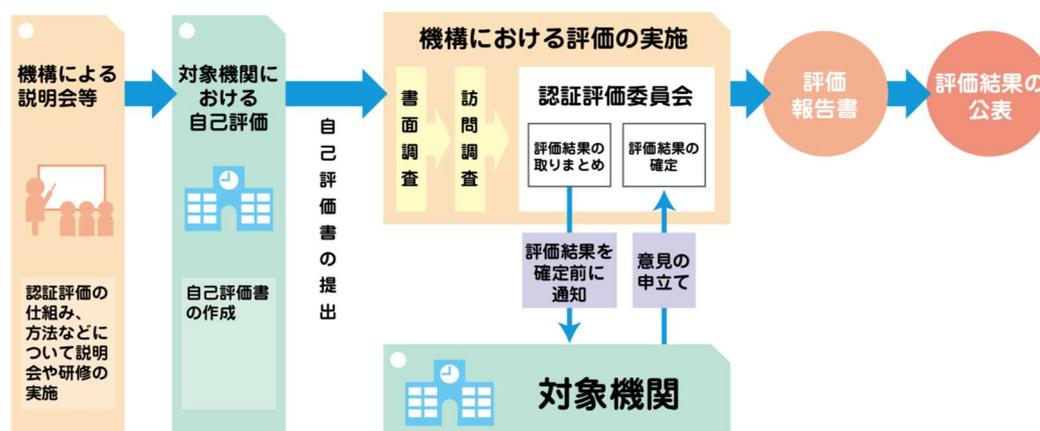


図 2-8 大学改革支援・学位授与機構による認証評価の流れ

出典：大学改革支援・学位授与機構概要 2017

<sup>14</sup> 大学基準協会，「大学基準」及びその解説，平成 29 年 3 月 21 日改定。

<sup>15</sup> 大学改革支援・学位授与機構，大学機関別認証評価 大学評価基準，平成 29 年 3 月改訂。

<sup>16</sup> 日本高等教育評価機構，大学機関別認証評価 実施大綱，平成 30 年度版(平成 29 年 4 月改訂)。

### 3. 研究開発評価の基礎知識<sup>17</sup>

本章では、研究開発評価を実施する上で必要となる基礎的な知識を提供するとともに、評価対象別にポイントを解説します。

#### 3.1 なぜ評価が必要なのか

そもそもなぜ評価が必要とされるのでしょうか。「大綱的指針」では、「評価は受動的なものではなく、その後の意思決定に活かしていくためのもの」であるとし、「あらゆる関係者は、評価の意義を再認識する必要がある」と意識改革の必要性が述べられています。逆に言えば、関係者の中で評価の意義が十分に理解されていないのが現状である、と言えるでしょう。

以下ではまず、評価の「関係者」とは誰なのかについて整理した後、いくつかの観点から、評価が求められる背景について解説します。

##### 3.1.1 研究開発評価の関係者とは誰か

研究開発評価の観点からみると、組織は次の3つの機能に分類できます。

まず1つ目は意思決定を行う査定部門であり、行政機関の場合、通常これは官房に集約されています。大学等而言えば、IR室など中期目標・計画の評価に関わる総務担当理事の所掌がこれに当たるでしょうか。

2つ目は、研究開発評価を担う部署です。こうした部署では、組織全体の研究開発評価体制や評価システムを整備するとともに、全体の評価結果をとりまとめたり新しい評価制度を導入するための様々な検討を行っており、評価システムの支援部署と位置付けられます。また、こうした部署では、組織全体にまたがる横断的な評価項目をもとに、例えば各事業担当部署で個別に展開された既存のプロジェクト群の中間評価や事後評価をまとめて実施する場合もあります。大学等にこのような組織は通常ありませんが、学内の研究力分析を行い、研究経営戦略のための意思決定を支援したり、競争的資金獲得やそのための評価対応を支援したりするURA室など研究担当理事が所掌する範囲にこれらが含まれるかもしれません。

これに対して、大部分の組織の個別事業担当課は所掌分野の事業を形成し、そしてまたそれを執行していくという役割を担う部署と考えられます。独立行政法人にせよ大学等にせよ、事業を推進する部署はすべてこれに該当するとも言えます。このような部署は事業を形成する際にも、またある制度のもとで研究開発プロジェクトを展開する際にも、自らが先ず事前評価を行わなくてはなりません。したがって、それぞれの担当部署は、評価の実施機能を備え、それを担うべき組織であり、評価スキルを十分に修得する必要があります。

このように、関連する機能の観点から評価についてみると、評価の支援機能と評価の実施機能、そしてさらに評価結果を踏まえて意思決定を行う査定機能という3つの機能に分けることができます。つまり、あらゆる主体では、組織としてこうした評価のための機能を備えていなければならない、また、組織の構成員は全員原理的に評価の当事者、関係者であると

<sup>17</sup> 本章は、財団法人政策科学研究所（現・公益財団法人未来工学研究所）「研究開発評価の質の向上のための調査・分析」（平成18年度文部科学省委託調査）の内容をベースに加筆修正を行ったものです。

も言えます。

### 3.1.2 環境の変化と組織経営

現在我が国が置かれている状況は、かつての「追いつけ追い越せ」というキャッチアップ体制の下で高度成長を遂げた状況とは全く異なる枠組みの中にあります。経済成長率でみると、我が国は 1973 年の第一次オイルショックの直後から低成長の時代に移り、また 80 年代の中頃には一人当たり GDP が世界のベストテン入りし、名実ともに経済的なフロントグループの一員にもなりました。それに加えて先進国全体の成熟社会への移行、社会全体の傾向としての知識社会化、少子高齢化の進展、90 年代に入ってから本格的なグローバル経済体制への移行など、さまざまな社会経済的環境が変化してきました。

このような環境の変化は、行政においてもかつての「行政管理」の視点からではなく、「行政経営」という新しい枠組みに適応した立場からの政策展開が必要になったことを意味します。5ヶ年計画を策定し、計画期間中は粛々と事業を推進するだけの時代から、事後チェックを実施し、政策の有効性を検証しなければならない時代に移ったと言えます。これは、行政という言葉で独立行政法人や大学等と言い換えても同様のことが成り立ちます。

また、低成長の時代では限られた物的・人的資源を有効に活用するために、資源配分において「選択と集中」が必要となり、その選択と集中のための論理（ロジック）を生み出すことが重要になります。実効的な意思決定ができるよう、現場の裁量権を拡大し、より機動的な経営ができるようにする必要があります。

フロントグループの一員としては、固有の問題を解決するために独自の戦略をつくる必要があります。その戦略形成機能が新たに必要となります。例えば大学のような教育機関においても、18 歳以下人口の減少に直面し、社会人教育等を含めた新たな経営戦略が必要となるでしょう。

このように組織をとりまく環境が変化したことにより、適切な経営戦略の策定、意思決定を行うために、評価を実施して情報やデータを集約・分析するということが必須の作業となります。

評価を実施することの第一義は、さまざまな情報を収集し、組織の経営・マネジメントに役立てるといった点にあります。

### 3.1.3 PDCA サイクルと評価

組織経営を実効的に行うためには目標管理システムが機能しなければなりません。製造プロセスを有するような企業では、目標を設定して活動を行い、それが効率的に行われているかをチェックするいわゆる PDCA サイクルが、ごく日常的に機能しています。PDCA サイクルとは、

1. 計画(Plan)を策定する
2. 計画にそって業務を実施する (Do)
3. 業務の実施が計画にそっているかどうかを確認 (Check) する
4. 実施と計画の双方を改善する案を検討し、その結果を次の計画に活かす (Act)

というプロセスのことです。この 4 段階を順次行って一周したら、最後の Act を次の PDCA サイクルにつなげ、螺旋を描くように一周ごとにサイクルを向上させて、継続的な

業務改善をしていきます。この業務改善のしくみをスパイラルアップ (spiral up) と呼び、ISO9000 (品質マネジメントシステム) や ISO14000 (環境マネジメントシステム) シリーズにも適用されている概念となっています。企業に限らず、あらゆる組織において PDCA サイクルは実施事業を効果的に展開するために重要かつ基本的なマネジメント・モデルと言えます。

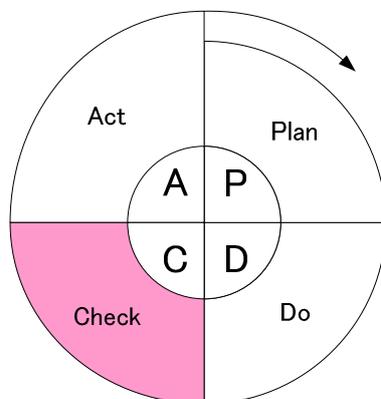


図 3-1 PDCA サイクル

出典: (財)政策科学研究所作成

PDCA サイクルが上手く機能するためには、Check と Act の部分が大変重要です。なぜならば、分析的観点から実施事業等を見ることで、スパイラルアップがなされるかどうかが決まるからです<sup>18</sup>。

評価は、評価対象である事業等の PDCA サイクルのうち、Check に相当するものであると一般的には考えられています。この機能が重要なのは、事業等を継続的に改善するあるいは次の事業に改善提案をするという Act のために位置づけられているからです。こうした機能が十分に発揮されるためには、Plan の段階で目標が達成された状態やそのために求められるアプローチ等が仮説として事後的に検証可能な形で設定されている必要があります。目標設定を含め、これらの仮説が妥当かどうかを検討するのが事前評価の役割です。また、Do の段階においても、達成状況等を把握するためのモニタリングの仕組みが必要です。その意味で、PDCA サイクルのあらゆる段階で「評価」が求められていると言えます。

#### コラム 2 : 評価を実施するとどのような効果があるのか？

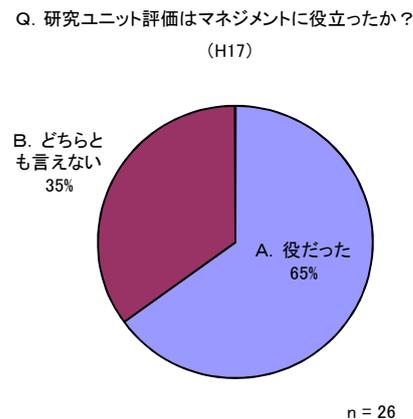
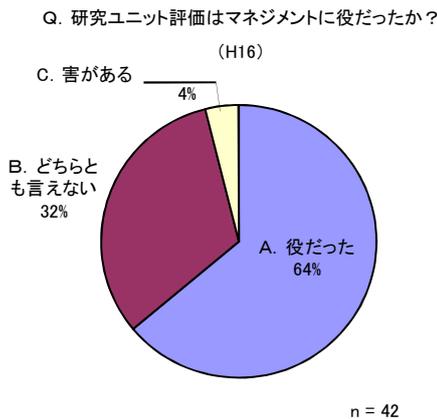
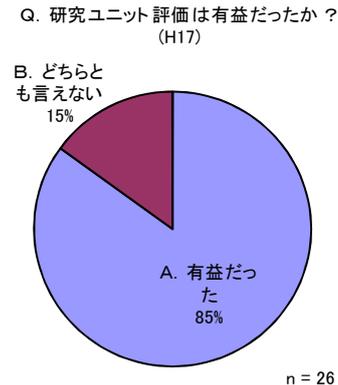
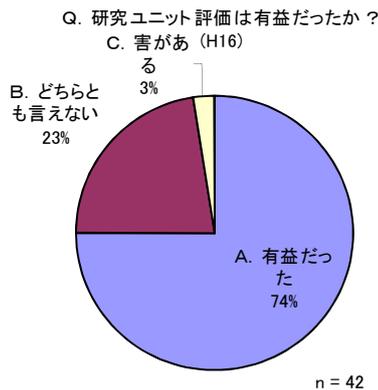
評価はマネジメントの改善のために実施すると解説しましたが、現実問題として、果たして本当に評価のおかげでマネジメントが改善されたかどうかを客観的なデータで示すことは難しいでしょう。しかしながら評価を導入して「有効であったかどうか」を主観的に判断することは可能です。

(独) 産業技術総合研究所 (現 国立研究開発法人産業技術総合研究所) では、第 3 期中期目標期間において研究テーマを関連づけて「研究ユニット」と呼ばれる研究組織を構築し、平成 14 年度からこれらに対して「ユニット評価」を実施していました。「ユニット評

<sup>18</sup> この 2 つの段階を統合し、PDS (Plan-Do-See) サイクルと呼ぶこともあります。

価」では外部レビューアから構成される評価委員会（レビューボード）が研究ユニットの実績を評価します。

古いデータになりますが、ユニット長に対するアンケート調査では、「研究ユニット評価は有益だったか」という質問に対して、平成 16 年度においても 74%が「有益だった」と回答しており、平成 17 年度ではさらにその比率が 85%に拡大しています。また、「研究ユニット評価はマネジメントに役だったか？」という質問に対しては、65%前後で「役だった」と回答しており、被評価者自身が評価の有用性を認識しています。組織運営において構成員の意識が前向きに変わることは非常に大切なことです。評価の効果はこのようなかたちでまず立ち現れるものと考えられます。



出典：独立行政法人産業技術総合研究所 HP<<http://unit.aist.go.jp/eval/ci/report.html>>

### 3.1.4 コンプライアンスとしての評価

現代においては、評価には組織におけるコンプライアンス（法令遵守）の側面があります。「コンプライアンス」とは、「法令や社会的モラルを守ること」ということになります。これは法治国家では当然のことなのですが、コンプライアンス違反の事例は民間企業や行政も含め、後を絶ちません。

コンプライアンス違反は社会的利益を無視して組織的利益を優先するために生じると言われています。企業であれば、「利益を出すため仕方なく」と言うでしょう。「同業他社でも同じことをやっているから」という理由で反社会的行為を行う企業、「売上を伸ばすため」と言う理由で法令を破る営業マン、「会社を守るため」と言う理由で粉飾決算等を行う経営

者などが該当します。行政や公的機関においても、組織防衛的な行動が予算の不正支出・流用等に結びつかないとは言い切れません。

したがって、どのような組織であれ、「法令を遵守し、モラルを保持し、社会に対しての説明責任を果たす」という側面がついてまわります。行政や公的機関であれば、納税者に対して、税金を使って事業をきちんと実施していることについて説明の義務があります。説明責任（アカウントビリティ）を超えて、積極的な情報公開と説明義務を担わなければなりません。

現在では評価業務を実施することは、前述のように、政策評価法や独立行政法人通則法、大綱的指針や各省の評価指針等で明記されているところです。今や、評価を実施することがコンプライアンスを確立し、社会的責任を果たすことに結びついていると認識しなければなりません。

### 3.2 評価が有用であるためには

以上、評価が必要とされる背景について見てきましたが、実効的な研究開発評価が行われ、その結果が適切に利用されるには、どのような仕組みが構築される必要があるでしょうか。評価を実施する際に常に心がけておかなければならない点は「評価する側と評価される側との間に一定の信頼関係が構築されていなければならない」ということです。そうでなければ、評価結果を組織的なマネジメントの改善に活かすことは難しいでしょう。

したがって、評価の枠組みを考える際に、査定的な態度が先に出てしまっただけではお互いの信頼関係の構築には至りません。どうすれば現状をより改善できるかという点を共通の目的に据えた上で、被評価者はデータや情報、意見等を提示し、評価者側はルールに則って評価するという枠組みを基本としなければなりません。

評価においてもっとも重要なことは、マネジメントの改善に役立てるために適切な評価の枠組みを構築し、関係者の理解と協力を得ることです。そのためには次のような評価の理念を確立することが有用でしょう。

#### 3.2.1 評価の理念を確立する

研究開発評価の理念とは、評価行為を成り立たせている根本的な原則のことです。前述のように、大綱的指針では、その冒頭に「研究開発評価の意義」が書かれてありますが、ここには、長い年月をかけて形成されてきたこの国の研究開発評価の理念なるものが凝縮されています。

##### (1) 研究開発評価の理念とは何か

研究開発評価は新しい問題に常に直面し、定型的な業務としてこなすことのできない実務的能力が試される仕事です。したがって、マニュアルだけでは対応できない状況において、基本に立ち戻って考える原理原則が必要になります。研究開発評価の理念を理解しておくことで、新しい問題に対してより有効な対処方法を考えることができるようになります。

研究開発評価の理念とは次のように言うこともできるでしょう。「評価の理念とは、評価の現場において個別具体的な問題に直面したときに、判断基準となる価値観、感性、美意識である。」

既存の理論に従ってロジックを積み上げ、それを現実に投射し、アクションを起こす、という合理性に基づく計画的思考は、さまざまな問題に直面する現場ではほとんど機能しません。より実務的な問題に対処するためには、行動の規範基準である理念が必要であり、関係者の中で共有されていることが大切です。

## (2) 研究開発評価の理念はどうあるべきか

研究開発評価のあるべき理念としては、「支援的であること (Supportive)」、「非排除・双方向的であること (Inclusive and Interactive)」、「明示的であること (Evincive)」の3つがあげられます。

この理念は、我が国の研究開発評価の現場の経験から導かれたものです。海外では異なる評価の理念が、独自に形成されています。その国の研究開発マネジメントや科学技術政策の仕組みに適切な評価の在り方が存在して当然ですし、文化的背景、歴史認識、政治手続によっても評価の理念は変わってくるでしょう。

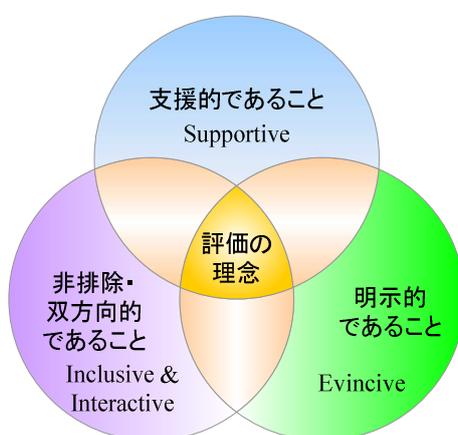


図 3-2 研究開発評価の理念

出典:財団法人政策科学研究所作成資料

### 1) 支援的であること

研究開発評価において支援的であることとは、研究開発実施者を「励まし、育む」評価であり、組織の長や政策立案者がよりよい政策・施策等を生み出し、進めることを促す評価であることを意味します。「励まし、育む」ためには最初から「切る」という姿勢では評価が成り立ちません。また、評価結果が査定と直接的に結びついているような仕組みも望ましくありません。なぜならば、評価を組織的な業務として定着させ、意義のあるものにするためには信頼関係の構築が最も重要になるからです。被評価者が評価実施側に対して信頼感を持つためには、評価を実施することが被評価者にとって有用であることが条件です。また、現実問題として評価を実施する場合に詳細な情報は被評価者側が持っているため、信頼関係の構築なくして正確な情報に基づく評価は無理であると言えます。

被評価者が評価を有用であると認識するためにはどのような措置が考えられるでしょうか。表 3-1 は、評価時期毎に想定される支援的措置を対象別にとりまとめたものです。

表 3-1 評価における支援的措置（例）

対象	事前評価	途上評価	事後評価
プロジェクト評価	<p>(プログラムの趣旨に照らして)</p> <p>【採択の場合】</p> <ul style="list-style-type: none"> <li>研究内容の改善に関する助言</li> <li>研究計画の改善に関する助言</li> <li>研究体制の改善に関する助言(他の研究グループとの協働等)</li> </ul> <p>【不採択の場合】</p> <ul style="list-style-type: none"> <li>不採択理由の提示(何が不足か?)</li> <li>改善提案(採択に結びつくには?)</li> <li>提案内容にふさわしい他のファンドの紹介、等</li> </ul>	<p>(プログラムの趣旨に照らして)</p> <ul style="list-style-type: none"> <li>研究内容・方法等の改善に関する助言</li> <li>研究活動に関する助言</li> <li>成果利用に関する助言(知財化、標準化等)</li> <li>リソースの追加的な割当、等</li> </ul>	<ul style="list-style-type: none"> <li>研究内容の発展に関する助言(次の段階のテーマは何か?)</li> <li>継続的な支援の可能性に関する提案(他のファンドの紹介等)</li> <li>成果利用に関する助言</li> <li>優れた研究に対する表彰、等</li> </ul>
プログラム評価	<ul style="list-style-type: none"> <li>より適切な目標への誘導</li> <li>指標の具体化や想定する受益者の特定に関する助言</li> <li>プログラムのポートフォリオに関する助言</li> <li>プログラム・マネジメントに関する助言</li> <li>プロジェクトの評価システムに関する助言、等</li> </ul>	<ul style="list-style-type: none"> <li>環境や受益者のニーズの変化を踏まえた目標の見直しに関する助言</li> <li>プログラム・ポートフォリオの見直しに関する助言</li> <li>プログラム・マネジメントの改善に関する助言</li> <li>プロジェクトの評価システムの改善に関する助言、等</li> </ul>	<ul style="list-style-type: none"> <li>次期プログラムの方向性に関する提案</li> <li>プログラム成果の展開に関する助言</li> <li>プログラム・マネジメントのあり方に関する教訓の導出</li> <li>プロジェクトの評価システムのあり方に関する教訓の導出、等</li> </ul>
機関(法人)評価	<p>【中(長)期目標・計画の策定】</p> <ul style="list-style-type: none"> <li>事業展開の方針や方向性、内容に関する助言</li> <li>具体的な目標の設定や測定方法に関する助言</li> <li>業務運営の改善や組織体制の強化・見直しに関する助言、等</li> </ul>	<p>【年度評価】</p> <ul style="list-style-type: none"> <li>事業展開の方向性や内容の見直しに関する助言、追加的支援措置の検討</li> <li>事業目標の見直し等の提案(上方・下方修正)</li> <li>業務運営の改善や組織体制の強化・見直しに関する助言、等</li> </ul>	<p>【見込評価】【期間実績評価】</p> <ul style="list-style-type: none"> <li>次期中(長)期目標期間における目標設定に関する助言</li> <li>次期中(長)期目標期間における事業展開の方針や方向性、内容に関する助言</li> <li>次期中(長)期目標期間における業務運営や組織体制のあり方に関する助言、等</li> </ul>

出典: 未来工学研究所作成

まず、プロジェクト評価、つまり研究開発実施者が提案した課題の事前(採択)評価において、評価実施側が応募課題の一定数を切らざるを得ない状況を想定してみましょう。採択された場合であれば、提案課題のより円滑な実施のために必要な措置を支持したり提案することが可能でしょう。しかしながら、不採択の場合においても、応募した研究者がよりよい課題を今後提案してくれるよう、「励まし、育む」必要があります。そのためには、なぜ不採択になったのか、どうすれば今後同様の課題に応募する時に採択されるようになるかについての情報を提示することが重要です。ファンドマネージャーが提案課題により相応しい他機関のファンドを紹介するというのも必要な措置として考えられます<sup>19</sup>。

途上評価の場合であれば、プロジェクトの途中段階でのコメントをレビュー評価から得ることができます。研究開発の最終的目標の確認、軌道修正や研究開発方法に対するコメントなど、研究開発実施側が気づかない点をフォローすることができます。実際のところ、我が国の研究開発プロジェクトでは途上評価時に研究開発の方向性や具体的目標が明確になるというケースが少なくありません。レビューアからの意見を参考にして、共通の目標を認識

<sup>19</sup>米国 NSF (National Science Foundation: 全米科学財団) 等の海外の資金配分機関では、プログラム・マネージャーが応募課題に対してより相応しいプログラムの紹介を実施するなどの工夫を行っています。

したり、研究グループ間の相互作用が生まれたりするという効能が途上評価には期待できません。なお、プロジェクト実施期間の特定時点で行われるこうした途上評価以外に、プログラム運営者がプロジェクトに対して日常的に行うモニタリングも重要です。プロジェクト実施者側と日常的にやりとりを行うことで、お互いの信頼関係やほどよい緊張関係を築くことにもつながります。

事後評価であれば「次にどうつなげるか？」が問題になります。研究開発テーマの発展性、継続性などに対してレビューアからコメントをもらい、支援的措置の可能性を研究者側に提案することができるでしょう。企業等が実施者である実用化研究の場合、事後評価は研究開発段階を脱して次の段階に移行できるかどうかを問う重要な評価になります。この場合も、研究開発以外の事業展開を含め、的確な支援措置を講じることが必要となります。

プログラム評価の場合、事前評価の段階では、より適切な目標へと誘導したり、ターゲットとする受益者をより明確にするといった必要性・位置づけの観点からの示唆であったり、より効果的、効率的なプログラムのたてつけやマネジメントのあり方に対する提案などを含んでいる必要があります。プログラム目標の達成に資するプロジェクト評価システムのあり方等に関する助言（よりよいプロジェクトを選定するための評価項目や基準等）もこの段階で求められることの1つです。

プログラム評価の途上評価では、当初想定していたプログラムをとりまく環境や受益者のニーズ等に変化がないかを確認し、必要に応じて目標の見直しを提言したり、ポートフォリオの観点から加えるべきプロジェクトのタイプを提案したりといったことが考えられます。事後評価では、実施期間中の成果を総括するだけでなく、当然ながら次期プログラムの方向性や成果の展開に関する提案などが期待されることになるでしょう。

機関（法人）評価の場合も、評定づけを行うこと以上の付加価値が必要です。たとえば、事前評価の段階では、中（長）期目標・中（長）期計画の策定に関して省庁のアドバイザー・ボード（国立研究開発法人審議会や国立大学法人評価委員会等）から適切な提案・助言が出せるかが問題となるでしょう。法人のミッションを具体的にどう展開し、目標とどのように位置づけるか、どの水準に目標を定めるかについての具体的な方針を協議する必要があります。これは年度評価や中（長）期目標期間における業務実績の評価の枠組みについて議論することにもつながりますので、機関評価の枠組みを支援的に構築する作業であると言っても良いでしょう。

年度評価（中間評価）では、詳細なデータに基づいた目標の達成見込みを踏まえながら、事業展開の方針に関する提案や支援的措置の必要性を検討することが重要になります。中（長）期目標期間における業務実績の評価（事後評価）では、次期目標期間における事業展開の方針についての前向きな提案が求められるでしょう。

プロジェクト評価にしる、プログラム評価にしる、機関評価にしる、被評価者側にとって評価が有用な機会でなければ、評価に必要な詳細なデータや情報を被評価者側が提出するインセンティブはありません。ましてや、最初から査定的な態度（つまり、「切る」評価）で評価を実施すると、お互いの信頼関係を構築することは難しく、継続的な評価の実施は困難になります。

## 2) 非排除・双方向的であること

研究開発評価において非排除的であるということは、被評価者が意見を述べる機会を排除

しないということです。研究開発評価では、一般的に外部評価者による評価パネルを組織し、ここでの検討が評価結果に反映されます。通常、評価は資料・データに基づくドキュメントを基に行われますが、評価者が被評価者の意見を直接聞くというプロセスも重視されます。この時に、被評価者である研究開発実施者は、ドキュメントには盛り込むことのできない研究開発への熱意、専門家としての展望、より広い研究の意味づけ等を評価者側に伝えることができます。

このようなやり取りが手続として原則的に確保されているということが、研究開発評価の枠組みの中で大切なこととなります。

### 3) 明示的であること

研究開発評価が明示的であることとは、評価のプロセスが明確な文書類に基づいて実施され、かつ評価結果が国民などに容易にアクセスしやすいかたちで公開されることを意味します。

評価が実施されるためにはその前提としてさまざまな文書規定が存在します。例えば、提案公募型の研究開発プログラムの場合では、プログラムの説明、公募要領、採択審査の基準が応募者に対して示されていなければなりません。応募者は公募要領に基づいて申請書を提出します。採択審査はそれらを基にメールレビュー等の手続を踏まえて審査されます。

採択審査の結果については、Web 等を通じて広く公開し、どの研究開発課題が採択されたかを明らかにします。また、不採択の課題については、申請者に対して不採択の理由書を送付します。

採択課題の提案者に対する契約においても、さまざまな取り決めが行われます。支出等の規定、評価時における情報開示義務、成果利用や知財化、上市の状況等の報告義務なども盛り込まれます。

研究開発評価実施時においても、成果報告書、ヒアリング記録、アンケート、それらをとりとまとめた評価報告書等の各種文書が作成されます。

このように、評価に限らず研究開発マネジメントで作成される各々の文書はお互いの記述内容を参照するといった一連のドキュメント・システムを構築しています。これは、ある文書が必ず他の文書の理由付けや根拠を述べているためです。最終的に評価結果を国民に公表するにしても、その結果に至った経緯を後から検証することができなければ意味がありません。

評価を含む研究開発マネジメントにおいて、明示的であることが要請されるのは、後から第三者が検証可能であることを担保するためです。

この明示性が確保されることで、支援的、非排他的であっても、馴れ合いではなく、緊張感のある関係をプログラム等の推進側と研究開発実施側とで保つことができます。

#### コラム 3：査定的な評価と支援的な評価

評価においては「切る」ための評価もあり得ると考える方も多いと思われます。例えば提案公募型プログラムにおける課題採択の評価は現実問題として応募者全員を採択するわけにはいかないのも、「切る」評価を実施せざるを得ないではないかとの批判もあるでしょう。

応募者に1回しかチャンスがないのであれば、その場合の評価は単に「切る」ことを目的

としていると考えても良いかもしれませんが、多くの場合は複数回公募を実施するでしょうし、同様の研究開発プログラムをその後も展開するのであれば、応募者は何回でもチャレンジできることが考えられます。

ここで重要なことは、提案公募型事業の推進者が応募者という潜在的な研究開発実施者を常にカスタマーとして認識しているかどうかという点です。事業がより良い展開を果たすためには、良質な課題が提案されることが条件であり、そのためには研究開発実施者が一定の水準でストックされていることが前提となります。事業が継続的に展開されていくためには、採択評価で「切る」にしても、研究開発実施者により良いリピーターとなってもらうための支援的な措置が必要です。例えば、不採択の明確な理由書、提案書に対する改善点の指示、他の資金配分プログラムの斡旋・紹介等が考えられるでしょう。

評価が「支援的であること」とは、評価を実施する側と被評価者とが継続的な関係を構築するための条件なのです。

### 3.2.2 評価対象を良く理解する

マネジメントの改善に資する評価を行うためには、評価対象をより良く理解しなければなりません。研究開発事業であれば、当該事業の研究分野における特徴、研究開発のリードタイム、失敗のリスクの大きさ等を踏まえ、適切な時期に適切な評価を行うことが重要になります。

これに対して、マネジメント側の都合であらゆる研究開発事業に対して一律の評価項目・評価基準を設定したり、異なる目的の研究開発に対して評価結果のみを数量的に比較したりすることは、評価そのものが事業のマネジメントと切り離されて実施されていると考えられます。

評価は基本的に一品料理であり、素材（評価対象）をいかに理解しているかが良い料理（＝良い評価）の決め手となります。また、和食の一品と洋食の一品とを単純に比較できないように、評価結果の単純な比較もできません。

### 3.2.3 形式的な評価にしないための工夫をする

日本では評価のプロセスをマネジメントサイクルに明示的に導入するよりも先に、各種評価制度が整えられてしまったことにより、評価業務がマネジメントと独立に存在するかのよような誤解が生じています。

行政機関の現場の声としてよく発せられるのが、業務のマニュアル化の必要性です。評価においてもマニュアル化を徹底し、要求される評価報告書を取りまとめ、現業の負担を少しでも軽くしようという考え方は当然の要請なのですが、マニュアルに従って評価を実施すれば良い評価ができるということにはなりません。

マニュアルでは最低限踏まえておかなければならない組織内の手順を示すことは出来ても、目的も背景も異なる様々な評価対象を適切に評価するための方法論を示すことはできません。せいぜい評価の観点を示すにとどまり、それを具体的に展開するのはマネジメント側（推進部署）の責任であると言えます。

また、マニュアルではあらゆる事態を想定して、一通りの業務が遂行されるよう構成されるので、実際に評価を担当する場合には当該事業にとって過度な手続が要求されることも考

えられます。評価を形式的に規定しようとするほど、多様な価値観を評価しなければならない研究開発のような事業ではかえって負担が増えることにつながりかねません。

評価はマネジメントの一部であるということを認識し、当該事業にふさわしい評価の在り方を事業の企画の段階で想定しておく必要があります。

### 3.3 評価の留意点

#### 3.3.1 評価の重複

現行の評価制度は評価を実施する側、受ける側から見ると必ずしも整合的であるとは言えません。研究開発評価は個別の研究開発事業等を対象としているのに対し、独立行政法人評価や国立大学法人評価などは、機関や組織（それを構成する研究グループ、研究者）を対象としています。

前者は事業等のマネジメントの改善に資することが主な評価目的となっていますが、後者は組織や個人に関する達成度評価の形式で評価されるか、場合によっては、他の組織・個人との比較（ランキング）まで行われることがあり、評価結果が査定等に利用される可能性があります。また、評価実施者もそれぞれ異なる場合が多く、前者の評価は現場に近い部署が担当し、後者の評価は組織の本部機能が担っているのが一般的です。

同じ研究活動を評価の対象としながら、評価の目的が異なるために、評価情報や評価結果を上手く活用できず、評価の二度手間や重複が発生しているのが現実です。

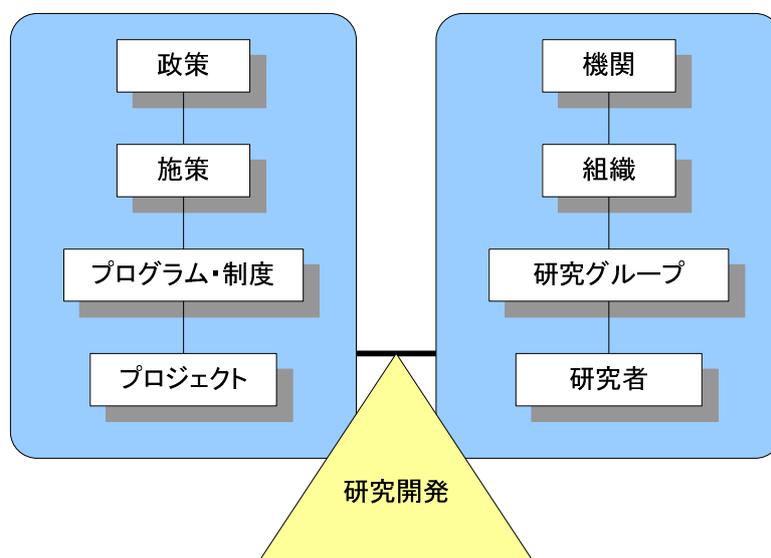


図 3-3 研究開発に対するさまざまな評価の局面

出典：財団法人政策科学研究所作成資料

しかしながら、こうした問題点は評価の基本的な枠組みよりも先に、評価制度が整えられてしまったことに起因していると言えます。本テキストでは、こうした問題点を踏まえ、どのようにすれば評価の重複が起らないように異なる評価制度に対応できるかについて考えてみたいと思います。

これから評価業務を担当する方々は、ある評価対象の評価に業務がとどまらず、他の組織や業務との連携が非常に重要である点を認識してください。

### 3.3.2 評価の費用

評価業務には一定の費用が掛かります。中間評価・事後評価を例にとると、直接的には外部パネルを招集するための経費（謝金、会議費等）、評価業務支援のための委託費などが相当します。また、評価業務に係わる職員の人件費もこれに上乗せされます。

一方で、目に見えないかたちで存在するのが、評価を実施するためにかかる内外の人件費です。事業推進部署の職員が評価業務のために資料を作成したり、研究開発実施者に対してヒアリングをするなどの費用や、研究開発実施者側においても必要書類を作成し、エビデンス類を確保するための費用が発生します。こうした業務を外注すれば委託費として一部は明らかになる場合もありますが、全てではないでしょう。

一般的に「評価に費用が掛かりすぎる」という認識があるとしたら、直接的な費用よりもむしろ目に見えない部分での負担が大きいという意味合いで語られることが多いかもしれません。

こうした費用に対する認識は、評価業務そのものの効率化によってある程度改善することも可能ですが、評価業務を本来のマネジメントの一環としてどこまで捉えるかということにも依存します。評価が文化として定着している組織においては、評価業務が負担であるという認識にはならず、やらなければならない仕事の一つとして内部化されています。むしろ、評価業務の量に見合った人員・体制・機能が確保できているかということが問題となるでしょう。

### 3.3.3 数量的評価のデメリット

評価結果を比較することが容易にできるため、数量的評価への期待は大きいものがあります。しかしながら、「どのような質をどのような数値で計測するか？」ということは非常に重要な問題であり、数量的評価を安易に適用するのは危険です。

例えば、研究者のアウトプットを評価する場合、海外査読付き論文数で評価したとします。このような数字が研究開発の予算配分や研究者個人の査定にまで利用されるような状況にあったとしたらどうなるでしょうか。

容易に想像できることですが、例えば70～80ページほどあるフルペーパーを専門誌に投稿せず、非常に短い分量のアカデミック・レターのような雑誌に投稿することで投稿論文の数を稼ぎ、それが実績評価値として採用され、さらにそのサイクルを助長することなどが考えられます。結果、研究の質を低下させることにもつながりかねません。こうした傾向は、任期内に成果を出さなければならない期限付き雇用の研究員の立場を考慮すると、さらに拍車がかかるものとなります。

では、論文の質が比較的反映されているとされるインパクト・ファクターを使えば良いでしょうか。論文が引用されるためにはある程度の時期を必要としますので、これも使用できる局面が限られてくるでしょう。また、使い方を間違えると、かえって問題を大きくしかねません（コラム4）。

数量的評価は、評価しようとする対象の一面を形式的に評価しているに過ぎません。それ

以外の面を捨象して評価することがないように、評価の枠組みを十分に検討する必要があります。評価結果が何かしらの査定に利用される場合には、特に注意が必要です。

むしろ、ミッションも活動の内容も異なる様々な機関・組織、あるいは研究領域の特徴や研究開発のリードタイムも異なるような研究者の評価を実施しようとする際に、比較可能な指標が容易に設定できるかどうかという点をよくよく考えることが重要です。

#### コラム 4：科学の「誤用」としてのインパクト・ファクターによる評価

インパクト・ファクター (impact factor、IF) とは、自然科学・社会科学分野の学術雑誌を対象に、その雑誌の影響度、引用された頻度を測る指標です。評価関係者なら、誰もが一度は耳にしたことのある言葉かと思います。

IF は、そもそも Clarivate Analytics 社 (当時は ISI 社) のデータベース Science Citation Index に収録する学術誌を選ぶ基準を作成するために、各学術誌の「影響」を示す指標として ISI 社の創設者であるガーフィールドらが開発したものです。彼は、機会あるたびに個々の論文の評価やその執筆者の研究能力の評価に IF を利用することは間違っていると指摘しており、その危険性は科学計量学や計量書誌学の研究者にとって常識となっています。

しかし、ざっとインターネットを検索してすぐにわかるように、IF が研究者の評価に使われる例は後を絶ちません。ある大学では、研究者として任期なし採用をするかを判断する際のクリティカルな要因の 1 つとなっていたりします。こうした使い方が間違っていることを示す一例として、調 (2009) は、ノーベル化学賞を受賞した田中耕一氏の例を取り上げています。同氏のノーベル賞受賞対象文献のうち、最初に発表されたのは IF が付与されない日中合同シンポジウムの要旨で、ついで初めて公表された論文が掲載された学術誌の IF は当時低く、2009 年時点でも 2.8 強でした。そして、対象となった最後の論文が掲載された雑誌にも IF は付与されていません。つまり、田中耕一氏はこの研究成果だけでは任期無し採用が果たせなかった蓋然性が高いと言えます。IF を使った評価はこのような暴力的な判断を行う可能性と常に隣り合わせにあることを意識する必要があるでしょう。

##### 参考文献

調麻佐志 (2009) , 科学計量学と評価, 科学技術社会論研究 (10), 16-28.

### 3.3.4 評価に対するインセンティブ

評価業務は評価実施者側と被評価者との協働で行われます。評価実施者側は多くの場合、それを業務として実施する必要性を持っていますが、被評価者は本業である研究開発事業の推進や研究開発に投じる時間数が削られるという認識を持つことが一般的です。

そこで、評価に対するインセンティブを十分に考慮しなければなりません。評価が望ましい形で実施されることが前提ですが、基本的には良い評価結果を受けたら研究開発予算等の配分に反映させるというように、インセンティブを考慮することが重要です。

逆に、誤ったインセンティブの枠組みを適用することによって、研究者等のモチベーションが深刻化する恐れがあります。

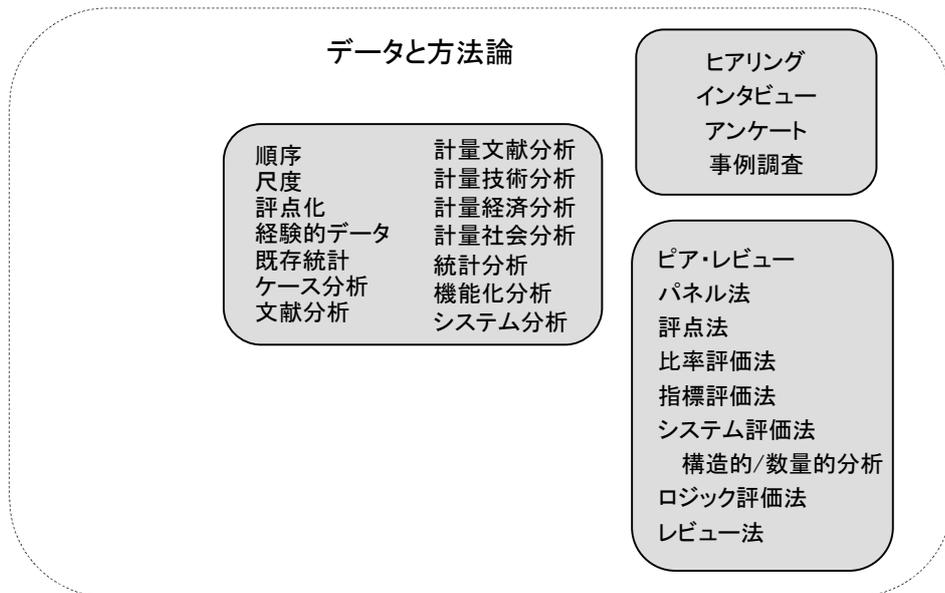
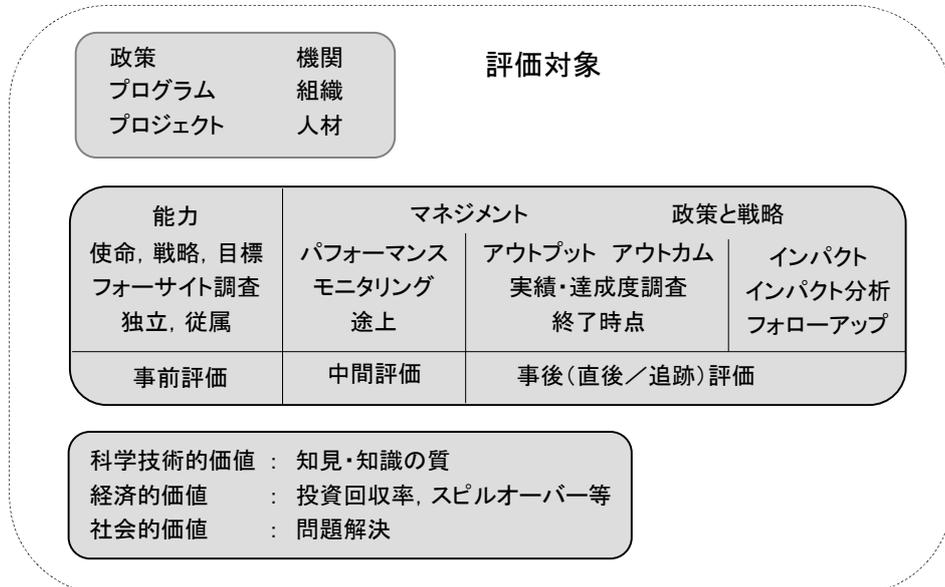
### 3.4 研究開発評価を適切に行うために学ぶべきこと

以上、組織で研究開発評価が行われ、またその結果が適切な形で利用されるための仕組みのあり方についてみてきました。以下では、実効的な評価を行う際に求められる知識について解説します。

必要な知識には様々なものがありますが、最低限理解しておくべきこととして、4種類の知識があります。1つ目は、前述の研究開発評価システム全体を貫く評価の理念であり、これがもっとも重要です。2つ目は評価対象をどのように捉えるのかです。3つ目は、研究開発の関わる複雑な対象について、実際に評価を行うためのデータをどのように測定・収集し分析するのか（メジャメントとメトリックス）、その結果をどのように評価に結びつけていくのかという研究開発評価の方法論です。4つ目としては、この対象の理解や評価法を担う専門家たち、つまり人材や評価作業の分担のあり方に関わるものがあります（図 3-4）。

## 評価の理念

支援的  
内在接触  
明示性



## 専門人材

運営人材 : 評価システムの設計・運営  
 評価人材 : ピア・レビュアー, メリット・レビュアー, 社会経済的レビュアー  
 アナリスト : 科学技術の側面, 社会経済の側面, その相互関係等の専門的分析

図 3-4 研究開発評価を行う上で必要な最低限の知識

出典: 研究開発プロジェクト等の評価手法に関する調査報告書(平成 13 年度経済産業省委託調査)

この他にも、研究開発評価を行う前提として身につけておくべき一連の知識があり、1) 評価一般に関わる考え方や知識（評価論）、2) 科学技術のマネジメントに関わる知識、3) 組織現象に関わる知識、4) 評価の運用に必要な基盤的な方法論に関わる知識などがあります（図 3-5）。

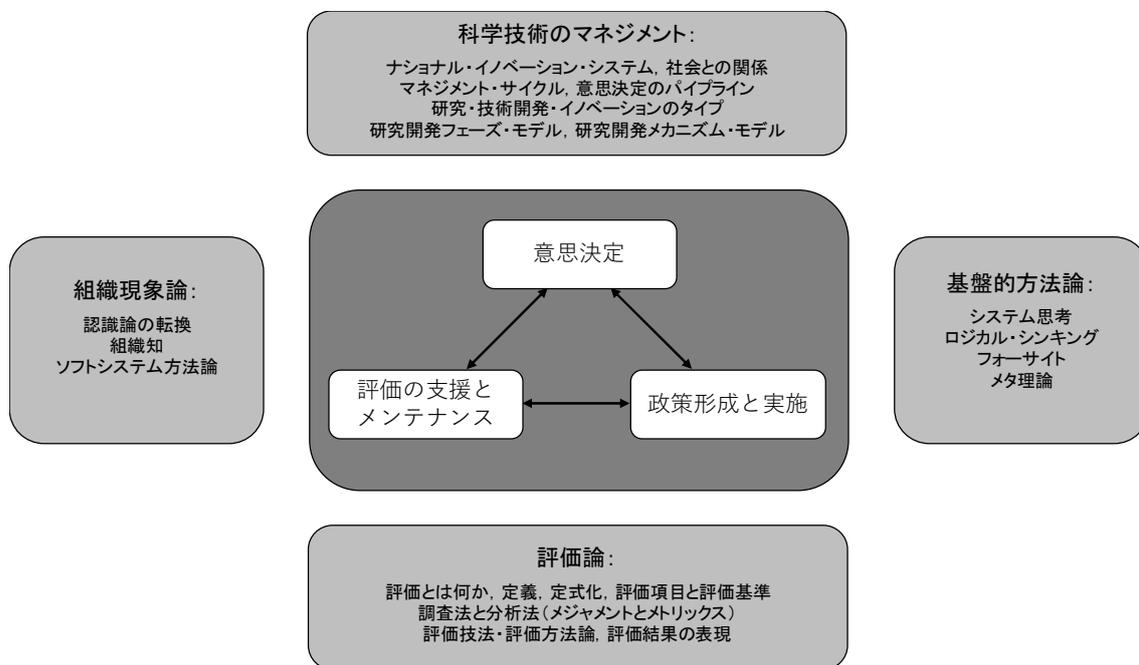


図 3-5 研究開発評価を行う上で身につけておくことが望ましい前提知識

出典: 研究開発プロジェクト等の評価手法に関する調査報告書(平成 13 年度経済産業省委託調査)

以下では、「研究開発評価を行う上で身につけておくべき最低限の知識」のうち、すでに詳述した評価の理念以外のものを中心に順を追って解説するとともに、「研究開発評価を行う上で身につけておくことが望ましい前提知識」についても「評価論」に関することを中心に必要に応じて適宜とりあげたいと思います。

### 3.4.1 評価対象の捉え方

研究開発評価において、評価対象はどのような側面から捉えることができるでしょうか。

ここでは、研究開発評価の **What** と **When** という側面から解説します。**What** とは、研究開発評価の対象についての理解や、研究開発事業等のいかなる実績を評価すべきかを表します。また、**When** は評価の時期（フェーズ）に関わるものであり、評価の時期によってどのような内容が重視されるべきかを見ていきます。

#### (1) 評価対象の区分

大綱的指針では、前述のように、研究開発評価の評価対象を 1) 研究開発プログラム、2) 研究開発課題、3) 研究者等の業績、4) 研究開発機関等の 4 つに区分しています。

各府省等が実施している研究開発事業等の制度的枠組みが異なるため、大綱的指針では上

のような区分を適用し、それぞれの評価対象についての留意点をとりまとめています。大綱的指針の区分と各府省等が実施している研究開発事業との対応を考えると混乱しやすいのは、大綱的指針の区分は原理的であり、実際の研究開発事業は事業形態に応じて区分されていることです。

大綱的指針の原理的区分とは、研究開発施策を政策ツールとしての枠、つまりプログラムとして捉え、研究開発課題をプログラムの下に位置づけられるプロジェクトとして捉えていることです。つまり、政策の階層構造を意識して区分されています。

一方、実際の各府省等が実施している研究開発事業は助成事業か、委託事業か、補助事業かという契約形態による区分、または、課題公募型事業か、委託先公募型事業か、もしくは特定の独立行政法人への委託事業かという区分にしたがって展開されており、基本的に事業を政策単位もしくは予算単位として運用しています。これは大学等や研究開発機関等においても同様のことが言えるでしょう。

研究開発評価において重要なことは、事業形態がどのようなものであれ、評価の目的に応じて評価対象を原理的に区別することです。これらの区分のうち、もっとも重要な単位がプログラムです。プログラムは組織や上位政策の目的と紐付けられ、複数のプログラムを組み合わせることでこれらの目的を達成しようとしています。したがって、機関評価を行う際もプログラム評価がベースになります。プロジェクト（研究開発課題）は、公募型か特定の主体への委託か等によらず、プログラム目的に資するかどうかという基準で選定されます。そのため、プログラムがきちんと設定されていないと、有効なプロジェクト評価は実行できません。逆に、プログラムがきちんと設定されていれば、プロジェクト評価の結果をプログラム評価に活用することも可能になります。このような原理的区分にしたがうことで、評価の重複を避け、評価を効率的に実施することが可能になります。

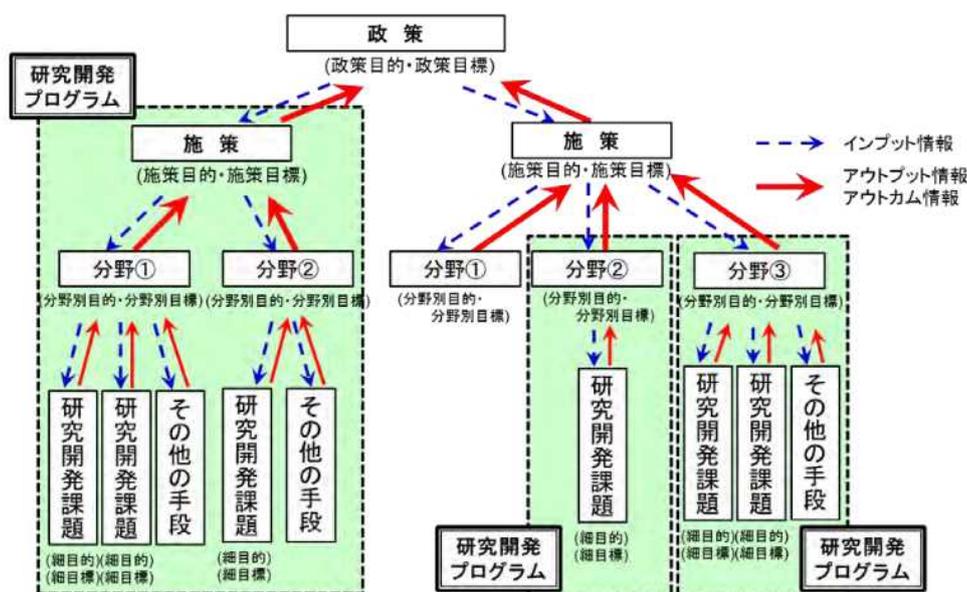


図 3-6 研究開発施策等の階層構造

出典：大綱的指針

## (2) 研究開発の実績の区分

評価項目のうち、もっとも重要なものは「実績」に関わる項目です。実績には通常、「何を産み出したか」に関わる「成果」の部分と、新しい取り組みとして「どのように産み出したか」に関わる「過程」の部分に分けることができます。

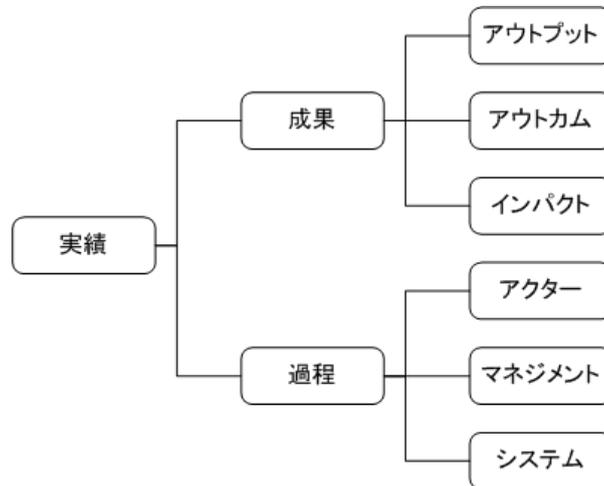


図 3-7 実績概念の区分

出典：財団法人政策科学研究所作成資料

成果を区分する概念としては、これまで「意図的／副次的」、「直接的／間接的」、「形式的／本質的」といったものが適用されてきました（表 3-2）。しかしながら、物事の多くは二元論で割り切れるものではなく、それらの区分概念が同時に組み合わせられて表現されるものと考えられます。

表 3-2 成果の区分

区分概念	定義
意図的 (intended)	目的・目標に沿った成果
副次的 (unintended)	目的・目標に含まれない成果
直接的 (direct)	事業等によって直接的に産み出された成果
間接的 (indirect)	事業等の成果の二次利用によって産み出された成果
形式的 (formative)	形式的(量的)に把握が可能な成果
本質的 (essential)	本質的な価値を有する成果、成果の質的側面

出典：財団法人政策科学研究所作成資料

こうした区分を適用すると、評価対象が上手く当てはまらない場合にどのように評価すべきかという問題に直面します。例えば、「副次的ではあるが、直接的かつ本質的な成果」、「意図的ではあるが、間接的な成果」という場合には、こうした類型化は成果の特徴を述べる分には有益かもしれませんが、評価のためのカテゴリーとしてはあまり意味を持ちません。

近年、「アウトカム評価」という言葉がよく聞かれるようになってきました。「そもそもアウトカムとは何か？」という議論よりも先に、なんとなく社会経済的価値に相当するものという理解が先に出来てしまい、当該事業のアウトカム探しをしなければならないという混乱が生じているように思われます。

ここでは、成果の区分概念として、基礎研究から実用化研究の場合においても適用可能な類型化概念を説明し、なぜ「アウトカム評価」が必要なのかについての理解を深めます。

## 1) アウトカム

最初に、難解なアウトカムの概念について学びましょう。

第2章で言及したように、大綱的指針では、「研究開発に係る活動自体やそのアウトプットによって、その受け手に、研究開発を実施または推進する主体が意図する範囲でもたらされる効果・効用」と定義されています。つまり、評価対象となる組織やプログラム、プロジェクトといったある活動の単位がどのような効果を上げたかを検証するための概念であり、研究開発に限らず、活動の意図（目的・目標）に依存して、アウトカムの内容や範囲は異なってきます。

ここで重要なことは、ある活動が目的・目標としてどのような範囲（時間的・空間的）で有効な手立てを考えているのか、ということです。つまり、事業等の責任範囲を明らかにすることが大切です。

### 目的と目標の設定

ここで、目的・目標とはどのようなものか整理しておきましょう。

目的とは「何のために（why）」行うかを説明するものであり、願望、期待などが反映された、現状とは異なる新たなステージのイメージであると考えて下さい。例えば、「CO<sub>2</sub>排出量を削減する」ことがある活動の主目的だとすれば、目的表現は「CO<sub>2</sub>削減に資する〇〇技術を開発する」といったものになります。

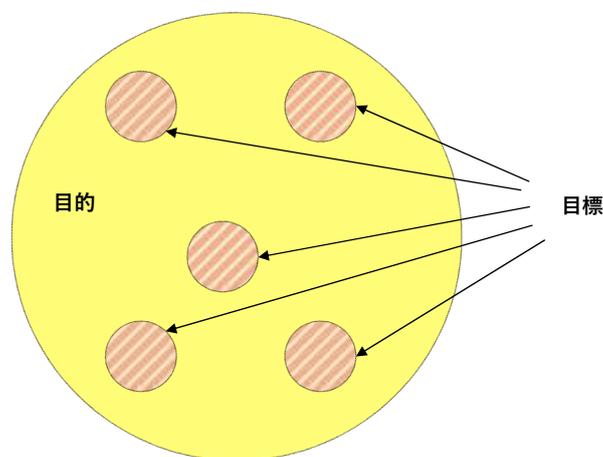


図 3-8 目的と目標

出典：財団法人政策科学研究所作成資料

目標の設定は、目的表現を具体的に観測可能な達成水準として示すことです。目標を定めるという行為は、茫漠とした目的の範囲をある程度限定し、もっとも実現すべき事項がどこにあるかという「境界」を定めることです。例えば、「CO<sub>2</sub>排出量を削減する」という目的に対して、「総量としてどれだけ削減するか?」、「いつまでに削減するか?」、「どの産業の

排出量を削減するか?」、「削減コストはいくらまで許容できるか?」といった具体的な目標が必要になります。

つまり、目標が設定されれば、その目標を達成することで活動の目的の全部ではないにしても、主要な部分が達成できることを意味します。

### 目的・目標とアウトカムの関係

目的・目標が明らかになったら、アウトカムをどのように考えるべきでしょうか。

プログラム等の活動は、それらを措置する何らかの理由に基づいて実現すべき内容が構想されます。これを活動の「目的」とすると、通常、「目的」は企画立案者の様々な願望を含んでいます。しかし実施に移される活動はそれが実現すべき内容を厳選し、確実に実現を目指す内容に絞られなければなりません。この厳選された内容を「アウトカム」とよび、活動の実施者が社会などに対してその実現を約束する、いわば契約内容を表しています。

したがって、アウトカムの対象・範囲は当該の活動を通じて実現を目指す「意図した結果」(intended results) に相当することになります。

では、目的・目標とアウトカムの関係はどのようになっているのでしょうか。下の図 3-9 は目的・目標とアウトカムの範囲を概念的に示したものです。「目的」はその内容を具体的に示す「目標」にブレイクダウンされますが、その際にアウトカムの内容(意図した結果)を代表するような目標を選びます。その目標の達成に向けて活動を展開することが、アウトカムをもたらすことになります。

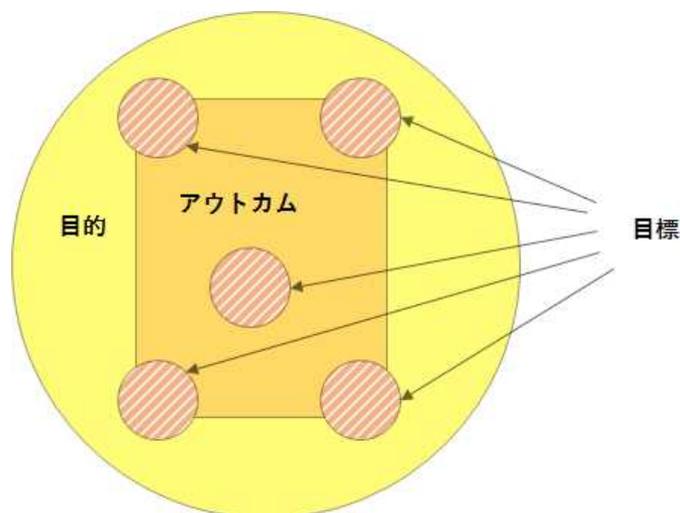


図 3-9 目的・目標とアウトカムの関係

出典:財団法人政策科学研究所作成資料

ここで重要なことは、アウトカムが具体的に何であるか、ということよりも、「目標の達成に向けて活動を展開すること」に注意を払わなければならないことです。プログラム等の活動が遠大な目的・目標を掲げつつ、実態は単なる試作品の開発や要素技術の開発だけに終わっている、というような研究開発の例は枚挙に暇がありません。目標として掲げた以上、それを実現するためにプログラム等としてどのような手だてを打っているのかが重要であり、その結果がどのような形で現れているかは別に検証すれば良いことなのです。

つまり、アウトカムが先にあるのではなく、活動の目的・目標とそれを実現するための具体的な措置があって、その結果（アウトカム）が何であるのかが問われる本質となります。

したがって、評価制度を変えただけでは、アウトカムは議論することはできません。プログラム等の企画・設計段階における目的・目標設定をきちんと展開するための、マネジメントツールが必要になってくるのです。

## 2) アウトプット

アウトカムが「意図した結果」であることを理解した上で、アウトプットとは何かを考えてみましょう。

大綱的指針では、「研究開発に係る活動の成果物。目的達成に向けた活動の水準を表す」と定義されていますが、一般的にはアウトプットとは入力に対して、あるメカニズムが産み出した出力結果を指します。プログラムの場合、その実施期間中にプログラムの企画立案者や関係する実施者が産み出したものがアウトプットということになります。しかしながら、「何を産み出したか？」という問いかけだけでは、アウトプットを列挙することはできません。アウトプットとは「目に見えて (tangible)」、「数えられるもの (countable)」であることが条件です。

下の図 3-10 は目的・目標、アウトカムとアウトプットの関係を示したものです。

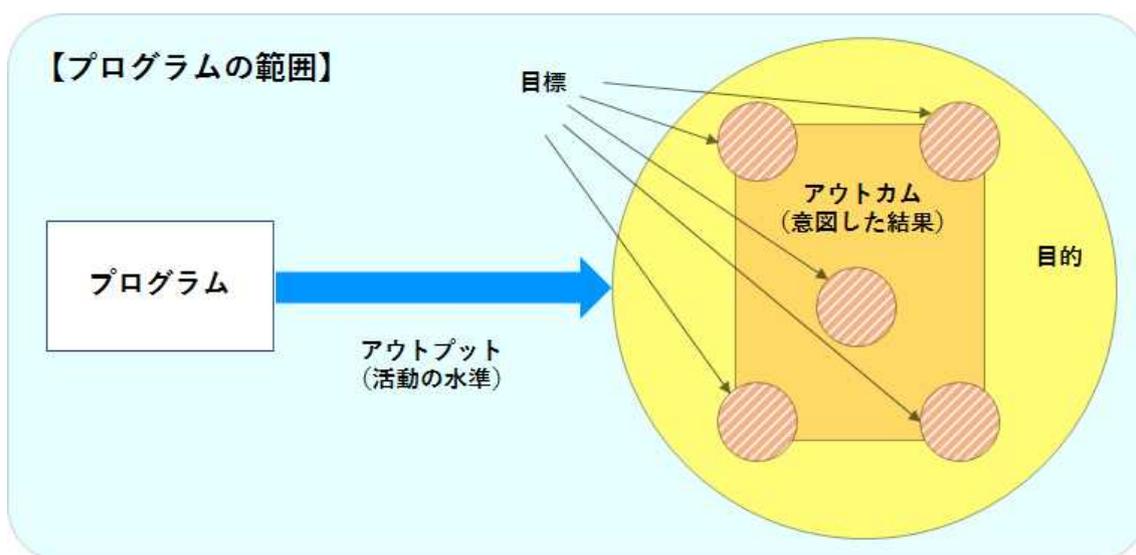


図 3-10 アウトカムからみたアウトプットの位置づけ

出典:平澤冷(2006), 内閣府総合科学技術会議発表資料

アウトカムが目的をブレイクダウンした目標の実現によって規定される「意図した結果」であるのに対し、アウトプットは目標の実現に向けた「活動の水準 (level of activities)」であると言えます。すべてのアウトプットは何らかの意思の下に実施された行為の結果として把握されます。その意思とは目標達成に向けた努力のことであり、その行為の結果とは、さまざまな活動の水準として把握されます。それらの水準を、目に見えるかたちに、数えられるものとして把握したものがアウトプットということになります。つまり、アウトプット

とは目標実現に向けた努力の証であり、できるだけ悉皆的に把握列挙する必要があります。

ここで注意しなければならないことは、アウトプットの数が多ければ多いほど良い、という論理にはならないことです。努力や活動の証ではあっても、価値判断はアウトカムの評価に委ねられます。アウトプットから何が産み出されたのかが問題であって、形式的に把握される数量に価値が見いだされることはありません。

例えば、論文の数が多ければ多いほど良いという議論にはならないはずですが、数が多ければ良いものが確率的に産み出されるという論法は、マネジメントを改善していくという考え方と相反するものです。アウトプットは単に活動の水準を表しているに過ぎないことを心に留めて下さい。「優れたアウトプットの創出を目指す」という記述が公的研究機関の評価指針などによく見受けられますが、本テキストの枠組みの理解からすれば、それはアウトカムを意味していることとなります。

### 3) インパクト

インパクトは大綱的指針において明確に定義されていませんが、アウトカムが意図的な結果であるのに対し、非意図的な結果全般を指すものとしてここでは考えます。下の図は先のアウトプット、アウトカムの関係図と関係者を対置したものです。「プログラムの企画者」とは企画立案者を指し、「プログラムの実施者」とは、プログラムの推進者や研究開発機関、研究者などの研究開発実施者ということになります。

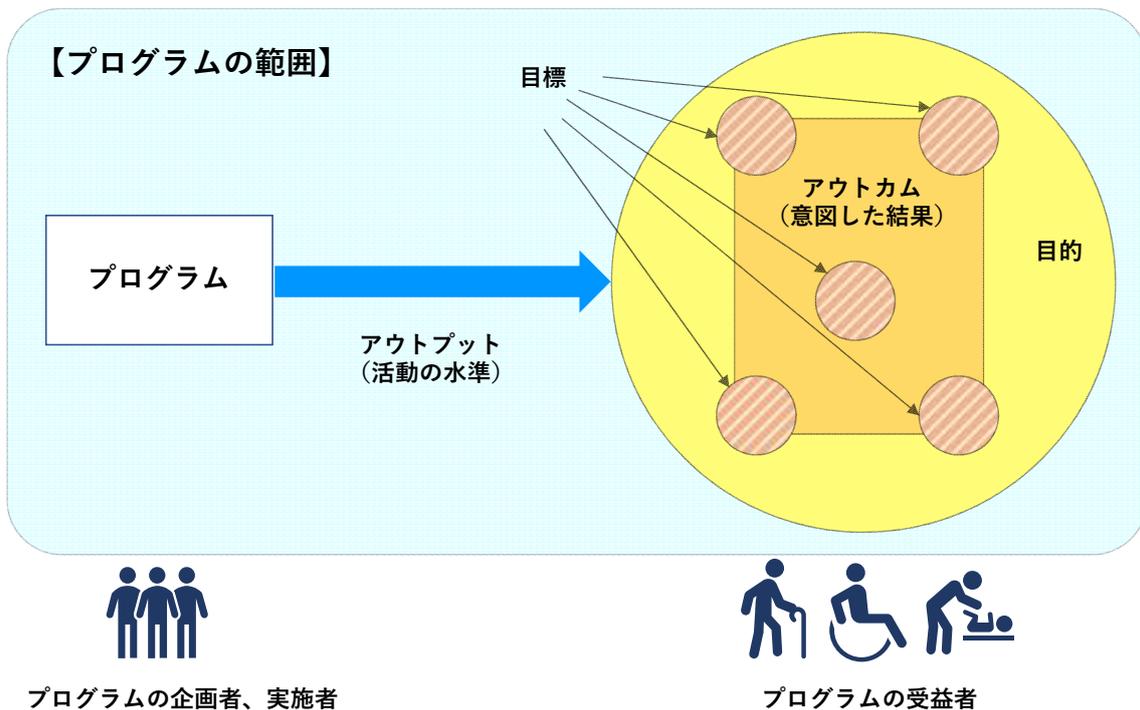


図 3-11 プログラムの企画者・実施者と受益者

出典:平澤冷(2006), 内閣府総合科学技術会議発表資料

今、何かしらのプログラムを展開したとしましょう。例えばある疾病の治療薬の開発を例

にとると、この研究開発プログラムが最終的に想定している「施策の受益者」は当該疾病に疾患している患者ということになります。

ここで、創薬を例にとると、新薬開発は前臨床段階と臨床試験の各フェーズをクリアしなければ最終的に薬が患者に提供されることはありません。ところが、大学や研究機関は有用物質の発見や化合物の生成について研究開発を実施し、治験の段階はもっぱら臨床を担当する大学病院等に委ねられることとなります。有用物質の発見や前臨床段階までの研究開発の成果は、薬理効果の可能性を示しているだけで、人体に投与して効果があり、かつ安全であるという保証を得るためには臨床段階を経なければなりません。

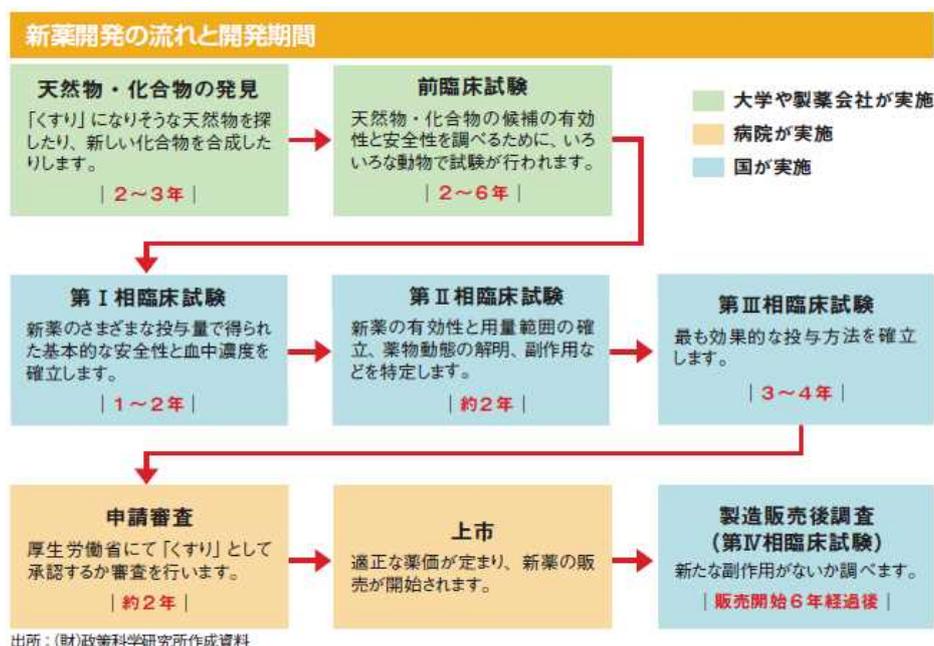


図 3-12 新薬開発のフロー

出典: (財)政策科学研究所作成資料

したがって、創薬に関する研究開発の場合、研究開発フェーズに応じたカスタマー（受益者）が誰なのか？という区別をすることと、研究開発プログラムそのものの責任範囲がどこまでなのか、という「線引き」が重要になります。前臨床段階までがプログラムの責任範囲であれば、特定の受益者（製薬会社や治験における対象患者）がプログラムの意図した最終受益者になるでしょう。臨床段階までを含み、医薬品市場における競争力の向上までを図るのであれば、特定の受益者からさらにすすみ、新薬の対象患者全体が最終的な受益者になるでしょう。

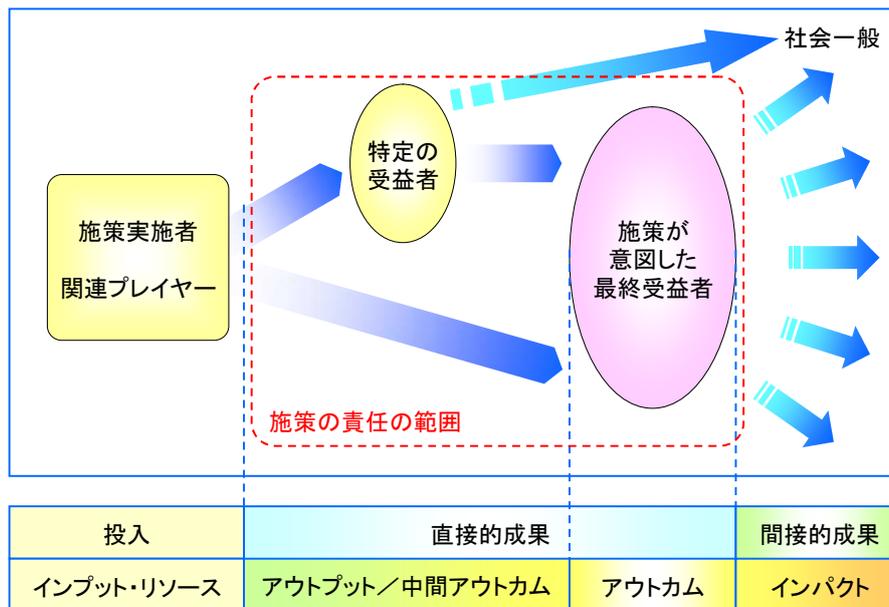


図 3-13 プログラムの責任の範囲とインパクト

出典:平澤冷(2006), 内閣府総合科学技術会議発表資料

アウトカムとインパクトを分ける判断基準としては、「プログラムとしてその効果を発現させる手だてを用意しているかどうか？」というのが妥当です。

創薬の場合、有用物質の発見が成果（アウトプット）であれば、薬としての本質的な価値（対象患者への薬理効果）がある程度確認された（前臨床試験が済んだ）時点でアウトカムが達成されていると判断できます。この場合のインパクトは、「有用物質から別の症状に効く新たな薬理効果が見いだされた」、「その発見を契機に有用物質の同族体がさまざまな研究開発機関で研究され、発見された」、あるいは、「それらの同族体の方が先に上市された」という場合も含み、プログラムの責任の範囲外で実施された行為によって得られた効果と判断することができます。

このことは、インパクトの大きさをもって研究開発の質を評価することはできても、インパクトをプログラムの実績そのものとして考えることはできない、ということを意味しています。なぜならばプログラムの誇大広告になってしまうからです。

#### 4) 成果以外の実績

研究開発評価では、成果以外の実績もきちんと評価しなければなりません。科学技術の振興のためには、優れた成果が創出されるように、人材育成や研究開発ネットワークの構築、研究開発環境の整備や地域クラスターの形成など、新しく効果的な仕組みを産み出したかどうかという、研究開発の「過程」についても評価することが重要です。過程では、主にアクター（実施体制）、マネジメント、システム（拠点形成などの仕組み）が評価されます。

基礎研究や戦略的な研究領域においては、評価実施時期に成果がかたちとなって現れるとは限りません。むしろ、政策の効果としては人材育成や研究開発拠点の形成、ネットワーク形成等において一定の効果が得られていることが重視されなければならないでしょう。

これらの過程における実績を定量的に示すことは、高度な分析ツールを必要としますが、

事業やプログラムの意図した効果がきちんと現れているかどうかを検証するには非常に有効です。

ここでは、分析ツールについての詳しい解説を省きますが、過程を評価する際の観点を概説します。

### アクター

研究開発プログラムや事業を展開する場合、さまざまなアクターが介在します。研究開発プログラムや事業の推進者（多くは官庁、資金配分機関）、研究開発の実施者（企業、研究開発機関や大学等の研究者）、成果の一次利用者、二次利用者などが想定できるでしょう。

それぞれのアクターにとって、研究開発プログラムや事業がどのような影響を及ぼしたかについて、明らかにすることが実績・実態把握の第一歩になります。

表 3-3 アクターに関する評価の視点

アクターの種類	明らかにすべき内容
研究開発実施者	<p>【人材育成面】</p> <ul style="list-style-type: none"> <li>-学位取得者数</li> <li>-受賞・表彰の有無</li> <li>-若手研究者・ポスドク等の参加度</li> <li>-男女共同参画の度合い</li> </ul> <p>【ネットワークの形成】</p> <ul style="list-style-type: none"> <li>-共著論文・共同発表の数</li> <li>-学術団体の分科会形成等</li> <li>-特許の共同申請状況</li> </ul> <p>【人材交流の実績】</p> <ul style="list-style-type: none"> <li>-国際共同研究の実績</li> </ul>
研究開発プログラムや事業の推進者	<p>【プログラムマネージャー(PM)の資質・配置状況】</p> <ul style="list-style-type: none"> <li>-専任のPMの人数</li> <li>-PMのバックグラウンド</li> </ul> <p>【マネジメント能力】</p> <ul style="list-style-type: none"> <li>-PMに対する研究実施者からの評価(アンケート等)</li> <li>-PMに対するプログラム・ディレクター(PD)の評価</li> </ul>
成果等の利用者	<p>【利用者の属性】</p> <ul style="list-style-type: none"> <li>-属性毎の利用者数</li> <li>-用途等</li> </ul>

出典：財団法人政策科学研究所作成資料

### マネジメント

研究開発評価において、マネジメントについて評価するといった場合、2つの意味が考えられます。

ひとつは、研究開発実施者側における R&D マネジメントの質についての意味、もうひとつ

つは、研究開発事業等の推進者側のマネジメントの質についての意味です。

研究開発実施者側の R&D マネジメントについては、巨額の研究開発プロジェクトや研究開発機関等の評価以外で、改善課題として問題にされることはあまりないでしょう。あまりにずさんな管理をしていた場合には、実施者側の管理能力が問われることは想像に難くないですが、このような例は多くの場合、契約事項や倫理規定への違反というかたちで対処ができます。

むしろ、公的研究開発資金を効率的に運用したかどうかの説明責任は、ファンディングを実施している側にあると言えます。すなわち、事業推進主体である省庁の原局・原課、資金配分機関の推進部署におけるマネジメントの在り方が問われることとなります。

また、マネジメント評価が対象とするのは、組織としての運用能力ですので、プロジェクトレベルよりも、プログラムレベル、機関レベルでの評価においてより重要になります。

ここでは研究開発の評価項目としてマネジメントを議論する場合には、研究開発実施者（研究者、研究チーム、企業等）のマネジメントではなく、ファンディング実施側のマネジメントと公的研究開発機関（独法、大学等）における組織的なマネジメントを中心的に扱います。

どのようなマネジメントが優れていたかの判断基準を明らかにすることは難しいでしょう。また、マネジメントが改善されたから優れた成果に結びついたということも証明することはなかなかできません。では、どのようにマネジメントは評価されるべきなのでしょう。民間企業のマネジメント評価のツールを参考にこの問題を考えてみましょう。

民間企業では、マネジメントの評価にはバランス・スコア・カード（BSC）がよく利用されています<sup>20</sup>。BSC では、「財務の視点（過去）」「顧客の視点（外部）」「内部業務プロセスの視点（内部）」「イノベーションと学習の視点（将来）」の“4 つの視点”からマネジメントの改善度を評価します。具体的にはそれぞれの視点を代表する指標群を設定し、改善目標をたて、それをモニタリングすることで業務改善の進捗度を把握します。

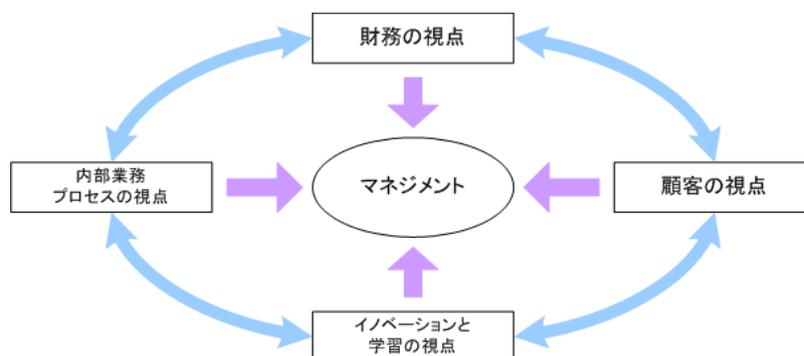


図 3-14 BSC によるマネジメントの評価軸

出典：財団法人政策科学研究所作成資料

財務の視点では、「財務的業績の向上のために、株主に対してどのように行動すべきか」ということを明らかにします。公的機関の場合であれば、「財務当局に対して、ひいては納税者に対して、どのような財務的改善の努力をするべきか」というように読み替えることができます。

<sup>20</sup> Robert S. Kaplan と David P. Norton らが提唱したマネジメント評価ツール。

顧客の視点では、「経営戦略を達成するために、顧客に対してどのように行動すべきか」ということを明らかにします。これも、「研究開発事業が想定している顧客（カスタマー）に対して、どのようにサービスの改善を図るべきか」というように読み替えることができるでしょう。

内部業務プロセスの視点は、「顧客と株主をより満足させるために、どのような業務改善を図るべきか」という点を検討します。つまり、業務の効率化指標を考えるということになります。

イノベーションと学習の視点とは、「経営戦略を達成するために、どのようにして変化と改善のできる能力や環境を維持するか」ということを検討します。これも、「優れた研究開発課題を発掘・支援するために、どのようにして変化と改善のできる人材能力と環境を維持・発展すべきか」というように読み替えることができるでしょう。

これらの視点を具体的な指標に落としたのが表 3-4 です。

表 3-4 BSC によるマネジメント指標の例

視点	評価指標(例)	単位	事業	機関	目標	結果
財務の視点	事業予算に対する人員数	円/人	○	△	対前年比(%)	
	事業予算に対する事業期間	円/日	○	△		
	事業予算に占めるGrantもしくは委託契約費の比率	%	○	△		
	事業費に占める人件費の比率	%	○	△		
	事業費に占める当期取得固定資産額の比率	%	○	△		
顧客の視点	普及・啓発費(全額)	円		○		
	新規応募数の伸び(対前年比)	%	○			
	クレーム発生件数	件	○	○		
	顧客からみた満足度(アンケート等による指標化)	-	○	○		
	Web閲覧数の伸び(対前年比)	%		○		
内部業務プロセス	内部業務日数の短縮	日	○	○		
	意思決定プロセスの明確化	-		○		
イノベーションと学習	担当人材一人あたりの研修日数	日		○		
	担当人材一人あたりのOFF-JT機会	回		○		
	担当人材一人あたりの研修費用	円/人		○		
	担当人材の満足度	-		○		
	担当人材の業務に関連した保有資格	件		○		

出典: 財団法人政策科学研究所作成資料

BSC を利用したマネジメントの評価では、PDCA サイクルを有効に機能させるマネジメント上の検討項目について、目標値を定め、それをクリアしたかどうかをチェックします。目標値が挑戦的であるかどうか、どの視点、検討項目を重視するかは経営戦略に従います。公的研究開発資金のマネジメントでは、経営戦略は事業の目的、あるいは機関のミッションに従うことになります。

マネジメントの評価は組織の自己改革に対する評価であるため、努力目標を自ら設定し、達成したかどうかを問うという、実務的なプロセスの繰り返しになります。マネジメントが功を奏したかどうかは、全体的な財務状況の改善や、組織に対する顧客の評価によって中長期的な指標として確認するしかありません。

## システム

研究開発事業やプログラムによっては、何かしらの直接的な成果を意図してではなく、優れた研究開発成果を生み出すための仕組みや、専門人材を育成するための制度を構築するという目的を持つものが少なくありません。

このような事業やプログラムにおいて、実績として評価すべき項目が「システム」です。

例えば文部科学省が以前実施していた戦略的研究拠点形成プログラムでは、「組織の長の優れた構想とリーダーシップにより、研究開発機関の組織運営改革や国際的に魅力ある研究拠点の創出を図る」とあります。

このような新しい仕組みを目的としたプログラムでは、その仕組み（システム）自体を評価しなければなりません。表 3-5 は戦略的研究拠点形成プログラムの中間評価における評価項目・評価基準を抜粋したものです。

表 3-5 戦略的研究拠点形成プログラムの中間評価における評価項目・評価基準

I. 進捗状況(目標達成度)(ミッションステートメントに対して)	
	<input type="radio"/> 計画に沿って順調に進捗しているか <input type="radio"/> 客観的な情勢の変化等(国立大学の法人化等)に対して適切に対処できているか <input type="radio"/> 当初の計画どおりに進捗していない場合、当初目的を達成する見込みはどれほどか
II. 組織運営の妥当性	
	<input type="radio"/> 調整費と調整費以外の外部資金、内部資金が、それぞれ組織運営構想に則って適切に充当されているか <input type="radio"/> 総括責任者は組織運営構想実現のための必要な権限と責任を有しているか <input type="radio"/> 単なる研究の実施だけになっていないか <input type="radio"/> 機関として組織運営構想の実現に向けて必要な取組・支援(人事制度・給与制度の改革等)を行っているか
III. 組織改革の成果	
(研究拠点としての波及効果)	
	<input type="radio"/> 組織改革により形成される研究拠点は、国際的な水準から見てトップレベルのものとなっているか <input type="radio"/> 組織運営構想の実施により、研究拠点としてのポテンシャルの増加が期待されるか
(組織運営構想の波及効果)	
	<input type="radio"/> 研究開発システム改革のモデルとなることが期待されるか <input type="radio"/> 他の研究機関にも波及し得る積極的な取組がなされているか
(情報発信)	
	<input type="radio"/> 広報は一般向けを含め十分に行われているか <input type="radio"/> 組織改革の取組について、関係機関への情報提供・情報交換により周知を図っているか
IV. 実施機関終了後における取組の継続性・発展性の見通し	
	<input type="radio"/> 実施機関の長による支援が十分行われてきたか、また、実施機関終了後も、それらの支援が継続され、機関本来の取組としての発展が期待できるか <input type="radio"/> 育成機関終了後においても、研究拠点となる体制を整備するための計画を有しているか <input type="radio"/> 研究拠点として機関終了後の発展性が期待できるか <input type="radio"/> 育成機関終了後の運営に必要な財政的・人力的措置が計画されているか

出典:文部科学省

システムの評価で重要な観点は、「事業等によって作り上げられた新しい仕組みが持続的なものであるかどうか」という点です。先の表の評価項目IVにあるように、事業が終了した後も継続して研究拠点が存続し得るかどうか問われています。初期の拠点形成のための投資が途絶えてしまったために研究拠点が存続できなくなれば、それは投資だったのではなく、単なる予算配分だったということになるからです。

### (3) 研究開発評価の実施時期と重視すべき内容

ここでは、いつ研究開発評価を実施すべきか(When)という評価時期の問題と、適切な評価時期にどのような内容を評価すべきかという問題について解説します。

## 1) 評価時期

研究開発評価の時期は評価対象である研究開発の時間的経過に則して決まります。評価時期は一般的に、「事前評価 (ex-ante)」、「中間評価 (mid-term)」、「事後評価 (ex-post)」の3つに区分されます。「事後評価」はさらに、「直後評価」と「追跡評価」に区分できます。

我が国の研究開発評価において「事後評価」と定義しているものの多くは、事業終了後1年以内に実施されていることから、「直後評価」に相当することになります。また、中間評価については、3年以上の研究開発期間を有する事業の中間年度に実施されることが一般的です。この中間評価以外に、日常的もしくは定期的にデータを収集し、進捗状況の把握を行ったりすることを「モニタリング (monitoring)」といいます。

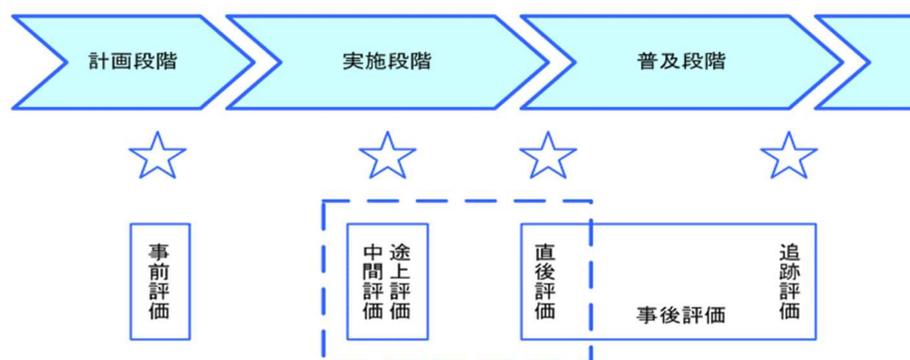


図 3-15 研究開発段階と評価実施時期

出典: 財団法人政策科学研究所作成資料

## 2) 評価時期毎の意義・重要性

研究開発評価の現場において、多く寄せられる疑問に「事後評価をなぜしなければならないのか?」という声があります。その理由は、中間評価であれば評価者のコメントを下に研究開発の軌道修正なり、改善提案が可能であるが、事後評価の段階では研究開発が終了しているので、前向きなコメントを得ることができない、あるいは評価結果を活かすことができないという意見です。また、さまざまな研究開発事業の終了年度が重なれば、事後評価を実施するだけで大変な負荷になり、その有用性もよく分からないというジレンマを評価現場では抱えています。

後者の負担感については評価の実施方法とも関連するので、ここでは取り上げませんが、事前評価、中間評価、事後評価、追跡評価が何のために実施され、どのような効用があるかについて、ここに取りまとめたいと思います。

### 事前評価の意義

政策評価法で義務づけられている以上に、事前評価の意義は大きいものがあります。なぜならば、後に続く中間評価、事後評価において、事前評価の内容が必ず参照されるからです。

研究開発事業が複数年度に渡って実施される場合には、事前評価の時期の担当者で中間・事後評価の時期の担当者が同一人物であるとは限りません。事業の実施理由と目的・目標が事前評価の段階で明らかにされていなければ、後任の担当者が困ることになるだけでなく、

中間評価や事後評価が事業に関する「そもそも論」で終始してしまう場合も想定されます。

事前評価において、事業の意義、目的・目標、期待される効果、事業推進者の所掌範囲が明確にされているからこそ、後に続く中間評価や事後評価において達成度評価が可能になるのです。研究開発評価において当面のわが国の課題は、事前評価をきっちりと実施するということです。

### **中間評価の意義**

おそらく研究開発評価を実施する上で、関係者の中で中間評価ほど有意義に捉えられているものはないでしょう。なぜならば、提案公募型事業や委託先公募事業では、事業の事前評価の段階で目的・目標に対する明確なコミットメントが出来ないために、中間評価の段階ではじめて関係者間における事業の方向性や成果の目標水準等の確認ができるという場合が少なくないからです。加えて、被評価者は、関連領域の専門家等から有益なコメントを得ることができる機会として積極的に捉えることができます。

研究開発の仕切り直しや事業推進側と研究開発実施者側が共通認識を形成する機会としても、中間評価は有用視されています。

### **事後評価の意義**

事後評価については、事業等の終了時点における目的・目標の達成状況についての確認もさることながら、次の事業等に活かすための知見を得ることが重要です。

また、事後評価は評価対象の品質保証（QA：Quality Assurance）を兼ねています。被評価者が研究開発実施者である場合、従事した研究開発実施者の実績にもなりますので、別の形での評価（組織内での実績評価、顧客の企業に対する評価等）に利用されることになります。事後評価でSを獲得したという品質保証が、研究開発実施者の次の研究開発への資金供与や融資等の機会に結びつくのです。

むしろ、事後評価の評価結果ほど本来戦略的に活用できるものはないと言っても過言ではないでしょう。評価が適切に実施されていることが大前提ですが、展開次第では、新しい研究開発案件の品質保証システムを構築することも可能だからです。

加えて、資金配分機関では、自らのマネジメントも評価項目として評価されます。事後評価において、資金配分機関のマネジメントが適切であったかどうかの評価は、機関評価のエビデンスとして活かされることになるでしょう。

研究開発実施者の立場から考えれば、事後評価が存在すること自体がある種のプレッシャーとなって、節度ある公的資金の使用とスケジュール管理を促す効果もあるといえます。

### **追跡評価の意義**

追跡評価については、フォローアップの観点が重要になります。政策の効果が果たして本当にあったのかどうかを確認するための作業といえるでしょう。したがって追跡評価については評価目的とする政策効果が十分に出現する時期を見越して、実施時期を設定しなければなりません。

追跡評価で確認する効果、特に社会経済的効果が、ある特定の事業の結果としてもたらされているとは考えられません。同様の社会経済的効果を目的とした事業を政策レベルで統合し、その政策としての効果が得られているかどうかを確認するというスタンスが評価の作業面からも有効となります。

#### (4) 評価時期毎の重視すべき評価項目

政策評価法や大綱的指針で定めるところの「必要性」、「有効性」、「効率性」の観点も、評価時期によっては、どのような評価基準で評価すべきかは本来異なります。このような点を踏まえ、評価時期毎に重視すべき評価項目を考えてみましょう。以下では主に国の事業を例に記載していますが、大学等や研究開発法人における評価においても、基本的には同様の考え方で捉えることができます。

##### 1) 事前評価で重視すべき評価項目

事前評価の段階では、研究開発はまだスタートしていません。ここでは研究開発の目的・目標設定や実施体制（責任体制）、研究開発計画、研究開発方法、コストなどについて、明示化された文章やプレゼンテーションをもとに評価します。

提案公募型の事業（プログラム）における個別課題（プロジェクト）の採択評価を除いて、一般に国が実施する研究開発事業の事前評価は政策評価当局（総務省）への提出資料である評価書に、他の事業と併せてとりまとめられます。これは、政策評価法の第十条に規定されている業務であり、各省庁とも政策評価担当部署（多くは官房）が推進部署からの提出書類を基に作成します。

事業評価書では、「必要性、効率性又は有効性の観点その他当該政策の特性に応じて必要な観点」を評価項目として記載する必要があります。この中で、事前評価の段階ではもっとも重要な評価項目は、「必要性」と「有効性」に関する記述です。

必要性とは、「なぜ事業を実施するのか（Why?）」が問われています。また、有効性とは、「その事業によって何を成すのか（What?）」が問われています。この二つについて詳しく見てみましょう。さらに効率性についても、事前評価の段階で記述すべき点を指摘しておきたいと思います。

#### 必要性

事業の必要性を述べるためにはどのように考えるべきでしょうか。直接的な答えは、「事業の正当性（rationale）」を論理的に展開することです。ある事業が必要であるという根拠として、形式的な側面と本質的な側面の両方から述べる必要があります。

形式的な側面とは、事業が実施主体の総合的な計画や政策の中でどのように位置づけられているかということです。これは、研究開発に限らず、政府の実施する事業は予算枠組み（一般会計か特別会計か）、政策体系の中で形式的に位置づけられ、国会等の承認を得てオーソライズされている必要があります。

本質的側面とは、形式的側面とは別に当該事業の意義を明らかにすることです。事業を行う前提条件としての問題認識・背景等が明らかにされ、対策の必要性を述べることとなります。また、国として実施する必要性を別途記述しなければなりません。

図 3-16 は必要性に関するロジックの展開例です。「事業の正当性」は、「政策の中での位置づけ」、「事業の意義」に大別され、それぞれがさらに論理展開されています。「政策の中での位置づけ」は、「政策体系の中での位置づけ」と「歴史的経緯の中での位置づけ」に分けられています。

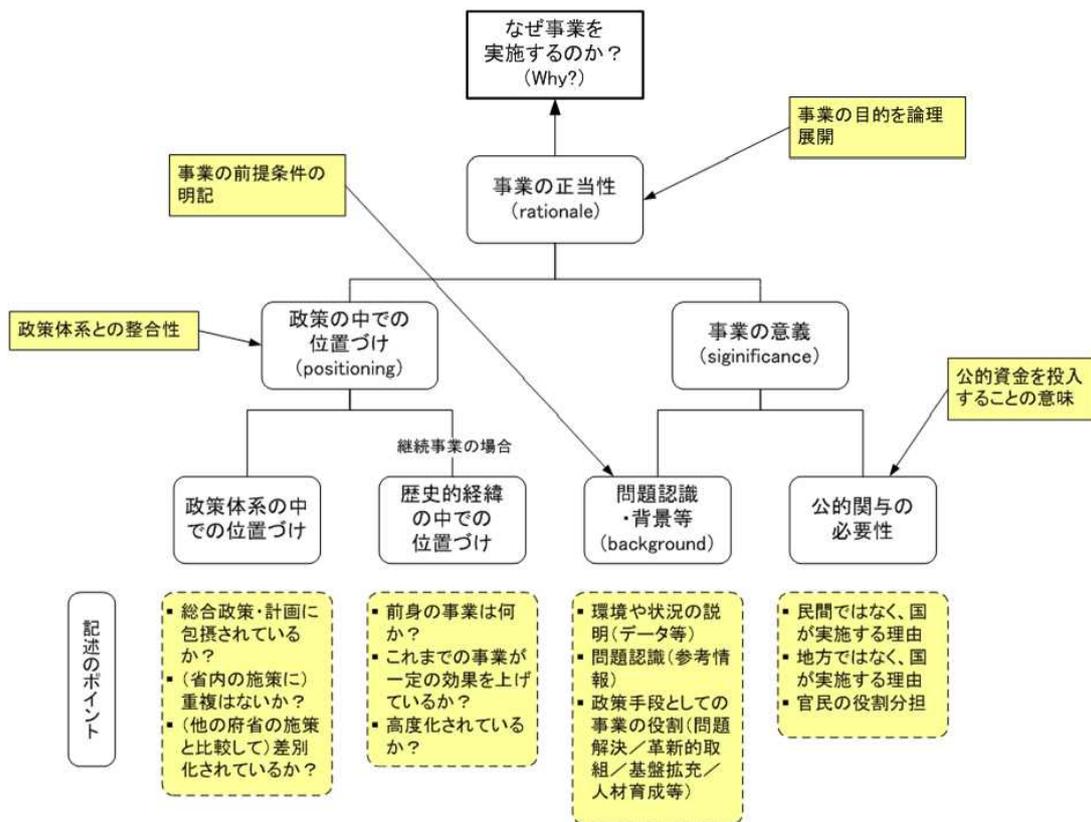


図 3-16 必要性のロジック展開（事前評価）

出典：財団法人政策科学研究所作成資料

「政策体系の中での位置づけ」は当該事業が、1)総合政策・計画に包摂されているか、2)（省内の政策の中で）重複はないか、3)（他省との政策と比べて）差別化されているかという点を述べることで網羅できます。

「歴史的経緯の中での位置づけ」は継続・拡充事業において特に明記しなければならない点です。1)前身の事業は何か、2)これまでの事業が一定の効果を上げているか、3)（これまでの事業と比較してどのように）高度化されているのかという点を明らかにします。

一方、「事業の意義」については、「問題認識・背景等」と「公的関与の必要性」に分けられます。「問題認識・背景」では、1)（事業が前提としている）環境や状況の説明、2)施策主体としての問題認識、3)政策手段としての事業の役割が明記される必要があります。

「公的関与の必要性」では、1)民間ではなく、国が実施する理由、2)地方ではなく、国が実施する理由、3)官民の役割分担等が明記されます。

「事業の正当性」を述べることは、事業の目的を裏付けるための記述に相当します。事業がフォーカスすべき目的がどこにあり、どのような役割が期待され、政策体系の中でどのように位置づけられているかを明記することで、実施事業の根拠が形成されます。

事前評価の段階で、必要性が明確に記述されていれば、中間評価や事後評価の段階で必要性を再度評価する必要がなくなります。もし、中間評価の時点で事業の前提条件に大きな変化があれば、改めて事業の意義を考える必要はありますが、研究開発に限らず、国の実施する事業は中長期的視野の下で展開されていることが一般的ですので、そうそう見直されるようであれば、政策評価のプロセスにおいて中止案件として議論の俎上に上がることになるで

しょう。

### **有効性**

事業の有効性を示すには、事業の目的と目標を設定し、そのために何を実施するかについて明らかにすることです。つまり、事前評価の段階で本来明らかにしておかなければならない点は、実施事業によって何が産み出され、どのような効果を国民生活にもたらすか、ということです。この段階で、中間評価、事後評価、ひいては追跡評価の際に検討されるアウトカムをどのように捉えているかが中心的な問題になります。

図 3-17 は有効性のロジック展開を目的と目標、アウトカムとアウトプットの観点から見ています。「事業によって何を成すのか？」という問いかけに直接的に答えるものは、「事業の目的」です。「事業の目的」では、事業の成果によってもたらされるであろう、国民生活や社会の新しいイメージ（ビジョン）が記述されます。これは、実現を確約するものではなく、政策担当者（部署）の願望・期待を込めた目的表現です。また、ここでの表現が、当然のことながら上位の施策や総合政策・計画等とリンクしていなければなりません。

当然のことながら、願望や期待だけでは困ります。事業実施主体として、何を実現させ、実現させるために何をするのかを明確にしなければなりません。ここで「事業の目標」を設定することになります。「事業の目標」とは、「事業の目的」で描いたビジョンを具体的に検討し、「これくらいの成果が得られれば事業目的に対して十分寄与した」と考えられる目標水準のことです。「事業の目標」は同時に、目標達成のための政策措置が及ぶ範囲を定める行為に他なりません。すなわち、事業の責任範囲であり、その結果はアウトカムに相当します。

「事業の目標」が設定されたら、その目標達成のために必要な活動を列挙します。事前評価の段階ではまだアウトプットが出ていませんが、事業目標の達成に資するアウトプットを想定し、そのための出力装置として事業の活動を設計することになります。

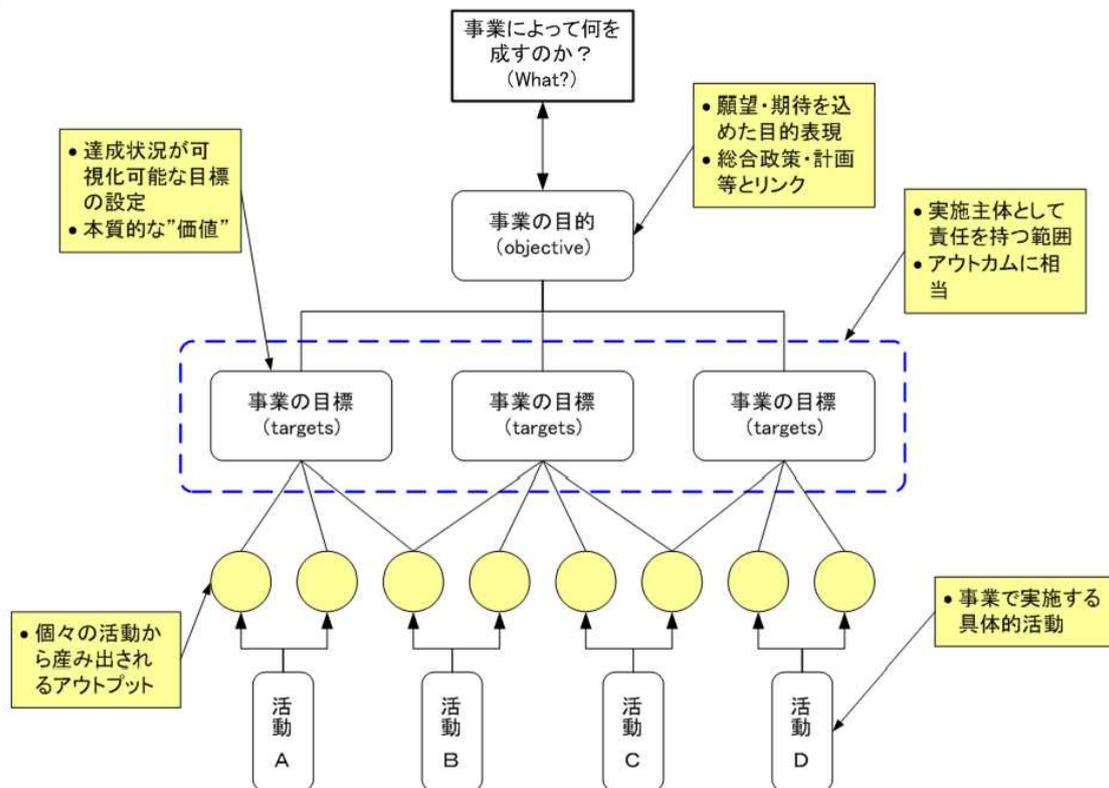


図 3-17 有効性のロジック展開(事前評価)

出典:財団法人政策科学研究所作成資料

### 効率性

事前評価の段階ではアウトプット、アウトカム等が得られていませんので、この評価の観点を厳密に記述することは不可能です。しかしながら、事業の個々の活動において、どのような体制・仕組み、予算等で実施するかについての記述は可能です。ただし、これらは事業を実施する段階において常に見直され、結果的に当初の体制や仕組み、予算配分で行われないことが想定されますが、中間評価や事後評価の段階で、そのようなマネジメントの変更が有効であったかどうかを検証することができます。

少なくとも、事業全体の管理機能、執行体制や責任体制については、事前評価の段階で明らかにしておくべき事項です。さらに、事前評価の段階では、事業全体のコストパフォーマンスについて、ある程度の比較考量を行う必要があります。他の研究開発手段と比較して、当該事業が目的達成のために優れた手段であることを裏付ける根拠を示すことで、当該事業の目的に対する効率性について述べることができます。

### 効率性記述の良い事例：

文部科学省平成 19 年度事業評価書（新規・拡充事業）より

「X線自由電子レーザー装置の利用開発」 研究振興局基礎基盤研究課

#### 【効率性の記述部分】

本装置は、大型放射光施設 Spring-8 に隣接して整備することで、世界で唯一、放射光施設と X 線自由電子レーザーが共存する施設となり、2つの光を同時に用いる実験がここでのみ可能となるほか、運用面でも効率化が実現される。

また、日米欧で熾烈な国際競争力が繰り広げられる中、我が国は高加速勾配加速管や真空封止型アンジェレーターなど独自に開発した技術により、欧米に比べ全長4分の1以下、総コスト2分の1程度と、コンパクト化と低コスト化を実現可能としている。

さらに、本事業のために開発した装置架台は、40年前に開発された送電線用絶縁体を利用することにより、従来用いられてきた架台の20倍という高い安定性を示しており、古い技術に新しい応用をもたらすなど、効率的な技術開発が進められている。

## 2) 中間評価で重視すべき評価項目

中間評価の段階では、事業の成果の一部が出されていることが想定されます。特に、提案公募型事業の場合、研究開発制度の下で採択された研究開発課題の一部は終了していることでしょう。

成果がある程度出現していることを前提に、中間評価において重視されるべき評価項目とは何でしょうか。再び、「必要性」、「有効性」、「効率性」の観点から、評価項目を検討してみましょう。

### 必要性

事業の正当性は事前評価の段階で確保されていると考えるのが妥当です。仮に、事前評価の時点と中間評価の時点とで、事業が依って立つ問題認識に大幅な修正を加える必要がある場合、改めて事業の意義を考えることとなります。

事前評価の段階で、事業が置かれている環境や問題の重要性等に関するデータを示しておき、データのフォローアップを行っていれば、中間評価時において事業の問題認識が妥当であるかどうかの確認となります。

### 有効性

中間評価の段階では、アウトプットがある程度出現しているため、「事業において何が行われたか？」を確認することができます。ふたたび、有効性のロジック展開図 3-18 を見てみましょう。中間評価で行わなければならないことは、事業の目標の達成状況を確認することです。既出のアウトプットが設定している事業目標に貢献しているかどうかをチェックされ、さらに事業全体で意図している結果（アウトカム）がもたらされそうかどうかを評価することとなります。

事業の目標達成状況が期待通り、もしくは期待を上回るものであれば、中間評価の評価パ

ネルの評点は高いものになります。一方、このままでは事業目標が達成できない、あるいは、アウトカム実現の観点からより良い研究開発方法、実施体制への示唆が得られたら、個々の事業の活動を見直すことになります。この場合、評点は低いものになるでしょう。

ただし、中間評価はあくまでも途中段階におけるチェックであり、仮に評点が低くても、中間評価時の評価パネルの意見を受け入れて、事業が改善されたのであれば、事後評価においては当該事業は高く評価されなければなりません。この意味でも、研究開発実施者や事業推進者にとって、中間評価は重要な意見交換の場であることが分かります。

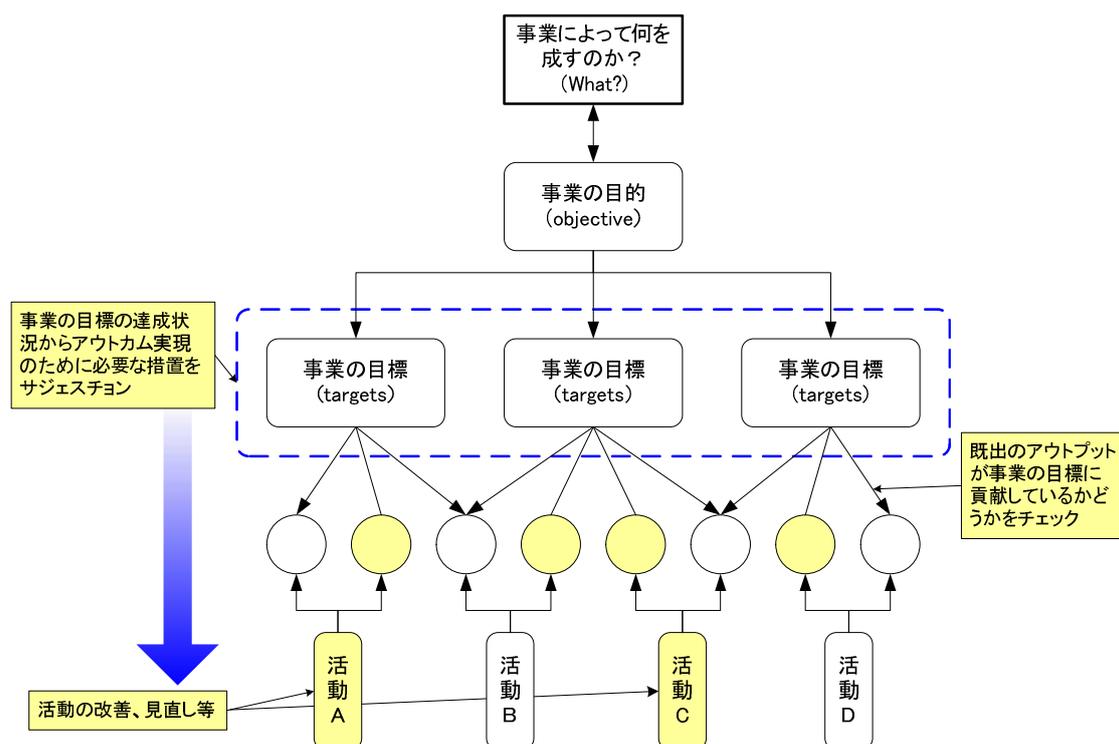


図 3-18 有効性のロジック展開(中間評価)

出典: 財団法人政策科学研究所作成資料

### 効率性

中間評価の段階では、アウトプットの発生状況が確認できます。これらのアウトプットを得る過程で、効率性の観点からチェックすべき点は、主に研究開発の実施体制の妥当性、研究開発実施側ならびに事業実施側のマネジメントの妥当性、資金ショートの場合の措置、研究開発方法等の見直しといった、主に事業運営に関わる項目となります。

効率性に関わる観点は、事業の目標達成状況と密接な関係があります。いくら研究開発体制やマネジメントが優れていると評価されても、肝心の成果が思わしくないのであれば評価した意味がありません。単純に、「マネジメントは妥当だったか?」と問いかけても、目標の達成状況が計画どおりであれば、「妥当である」としか答えようがありません。

したがって、中間評価の段階で効率性の観点から評価する場合には、どのようにすれば事業運営をより良く改善できるかという観点から実施されるべきです。

例えば、提案公募型事業の場合、提案書申請に関わる手続の効率化、採択審査プロセスの改良等について制度実施期間中にどのような試みを行った結果、課題の応募数が増加してい

る等の裏付けが得られれば、マネジメントの改善を明確に示すことができます。

### 3) 事後評価で重視すべき評価項目

事後評価の段階では、事業終了に伴い、すべてのアウトプットが出揃っている状況にあります。事後評価の段階では、研究開発実施者に対して研究開発の軌道修正や改善提案を行う必要性がないために、その効用を疑問視する声もあります。

評価時期毎の意義の部分で解説したように、事後評価の一つの目的は当該事業の品質保証を行うことにあります。公的資金を投入した研究開発事業が、当初の目標を達成し、かつ、質的に優れていたかどうかを検証する場が事後評価と言えます。

もう一つの目的は、事業実施主体のマネジメントに関わる問題です。当初の目標設定が目的に対して妥当であったか、研究開発体制として不備はなかったか、資金量は適切であったかに関わる事後チェックとして機能し、後継の事業や、新規の事業展開のための教訓を得る場となります。

ここでも、「必要性」、「有効性」、「効率性」の三つの観点から見てみましょう。

#### 必要性

事後評価において事業の必要性を評価するという事は、事業の目的とするものが有意義であったかどうかを判断することです。

事後評価において事業の意義を再確認するという行為は、仮に有効性、効率性の観点から当該事業が低い評価を受けたとしても、「事業目的に対して目標設定や研究開発方法、マネジメントなどが適切でなかった」という問題点の特定化のために必要なことと言えます。

また、事前評価の段階での状況と、事業が終了した段階での状況とを比較して、依然として事業の意義が失われていないのであれば、実施機関として次の施策展開を考える必要があります。

#### 有効性

事後評価においては、有効性の評価がもっとも重要になります。事後評価の段階で有効性を評価するという事は、事業の目標の達成度を評価することです。再び、有効性のロジック展開図を見てみましょう。

まず、当該事業におけるアウトプットを悉皆的にリストアップすることから始めます。このとき、事業として意図していない副次的な成果も含めます。

次に、それらのアウトプットが事業の目標に対して寄与しているかどうかを精査します。ここで、事業の目標の達成状況が確認されます。

事業の目標の達成状況が確認されたら、事業の目的、すなわちアウトカム(意図した結果)が実現しているか、あるいは実現しそうかについての評価を行います。

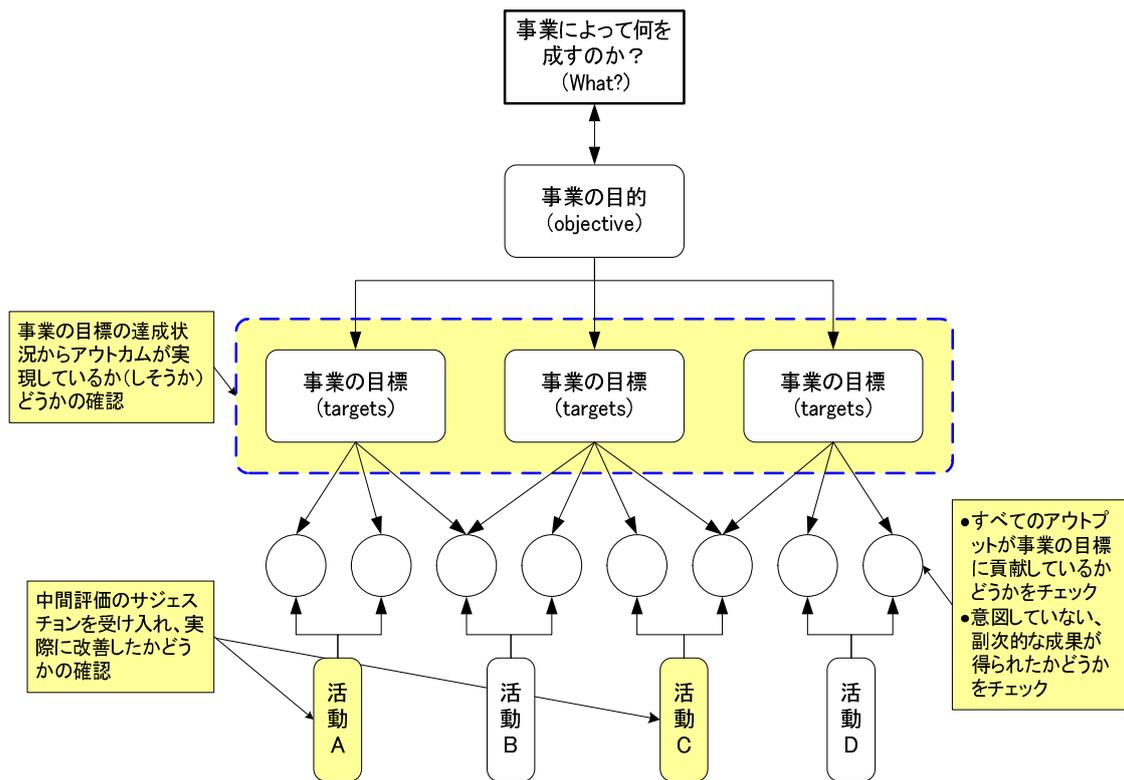


図 3-19 有効性のロジック展開(事後評価)

出典: 財団法人政策科学研究所作成資料

### 効率性

事後評価の段階ではすべてのアウトプットが出揃い、事業の目標達成度が明らかになります。この段階において、はじめて社会経済的な効果について検討することが可能になります。事業の目標がそのような効果に対して設定されていたとすれば、この段階において効果の確認を必要とします。

しかしながら、市場導入や普及を伴うような効果の計測については、その可能性に対する精査が必要になります。市場関係者の意見聴取、開発された技術や試作品の引き合い等の実績がなければ、多くの場合は絵に描いた餅となりかねません。

基盤技術、観測技術、計算科学技術等における開発の成果については、その後の利用計画を明確にすることで、社会経済的な効果を見積もることができます。

マネジメント等の評価項目に関しては、中間評価時の評価パネルからの提案を受け入れ、改善努力をしている場合には、高く評価する必要があります。

### 4) 追跡評価で重視すべき評価項目

追跡評価では、事業終了から一定期間経過後、事業の目的が達成されているかどうかをフォローアップすることがもっとも重要になります。ここで、事業の目的とはビジョンや願望を含めた、「社会の新しいステージ」を意図したものです。

一般に、アウトカム評価では、ターゲットとする社会経済的な効果を指標のかたちで限定し、その変化を見ることで、政策の効果を勘案します。厳密な因果関係はともかくとして、結果がよければ、施策として展開した各事業が意味を持っていたと考える検証の仕方です。

例えば、ある技術開発プログラムの目的が、「技術開発を通じた新市場の創出と、企業競争力の強化」といったものであった場合、プログラムに参画した企業の製品開発状況、市場化の状況、収益の状況等の情報を収集し、プログラムに参画していない同業他社との数値と比較するなどします。この結果によって、一定の効果が認められるようであれば、プログラムは意味のあるものであったと解釈することができます。

科学技術の場合であれば、科学技術領域別の論文被引用度等のさまざまな指標に関する国際ランキングをモニタリングし、わが国の科学技術政策の効果を論じることが可能です。

追跡評価で得られるデータはどのようなものであれ、今後の政策研究に必要なデータとなります。いかなる政策も単純に失敗だったということは断定できないでしょう。国等が展開する政策は何かしらの影響を持って、現実の社会を構成しているからです。しかしながら、どの部分は効果的であり、どの部分では無効だったという点を整理し、次の政策につなげるよう、データを収集することは喫緊の課題です。「思いつき」レベルの政策が「失敗だった」と言われないう、政策自体の高度化のために追跡評価は必要であると言えます。

### 3.4.2 研究開発評価のデータと方法論

ここでは研究開発評価の方法論（How）についての概要を学びます。はじめに研究開発評価の方法論の概要を把握し、次に効果計測の考え方や指標の見方について解説します。さらに、研究開発の質に関わる評価方法として一般的なレビュー法の基本について解説します。

#### (1) 方法論の一般論

##### 1) 調査＝分析＝評価のための方法論と意思決定

研究開発評価を実施するためには、評価対象に関する情報やデータをまず収集し、これらを分析した上で、評価の俎上に載せる必要があります。正しい分析のためには綿密な調査が必要であり、正確な評価のためには正しい分析が必要であるという意味で、これらのプロセスは互いに密接に関係しています。しかしながら、このプロセスは機能としては独立であり、研究開発評価の業務プロセスを考えた場合には区別しなければなりません。

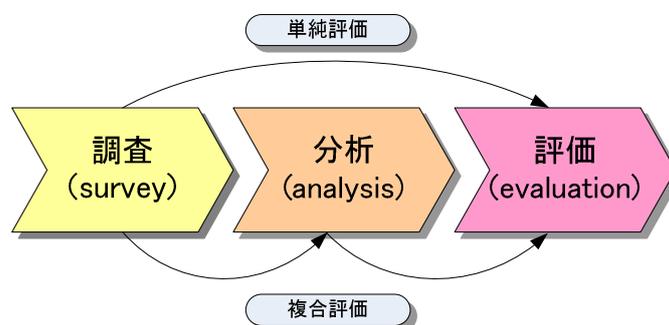


図 3-20 研究開発評価のプロセス

出典：財団法人政策科学研究所作成資料

特に、研究開発に伴って得られる情報やデータのみによ拠し、評価の目的を構造化して専

門的な分析することをせず、レビューに諮るような評価の仕方を「単純評価」と呼びます。我が国の例で言えば、科学研究費助成事業（科研費）の個別課題に対する評価がこの方法であると云えます。

一方、単純評価で使用される情報やデータ以外に幅広く情報を収集し、評価の専門家（アナリスト）がその専門性を発揮して分析を加え、その結果を下にレビューに諮るような評価の仕方を「複合評価」と呼びます。機関評価や実施機関側の何かしらの政策目的・ミッションに基づいて行われるプログラムの評価においては、効果計測・分析のプロセスが必要になりますので、複合評価でなければ評価したことにはなりません。

表 3-6 は調査＝分析＝評価のための方法論を体系的に示したものです。

表 3-6 研究科発評価の方法論

調査法		分析法	手法区分	評価手法
情報収集	データ収集			
インタビュー(関係者) ヒアリング(専門家) アンケート調査 事例調査 文献調査 顧客調査 社会調査	既存統計 アンケート調査 申請・管理データ データマイニング テキストマイニング : : :	ケース分析 コンテンツ分析 計量文献分析(ビブリオメトリクス) 引用分析、共引用分析、共語分析 謝辞分析 計量技術分析(テクノメトリクス) 特許分析、サイエンスリンケージ分析、 技術仕様分析 計量経済分析(エコノメトリクス) 費用分析、費用便益分析、費用効果分析 生産関数分析、消費者余剰分析 財務指標分析 計量社会分析(ソシオメトリクス) 表明選好法、各種指標分析 多変量解析 システム分析	定性的評価	レビュー法 メールレビュー、ブラインドレビュー パネル法 ピアパネル、エキスパートパネル ボードパネル
			半定量的 評価	評点法
			定量的評価	指標法 単一指標、複合指標 比較評価法 対比較年度比、対計画比 コントロールグループ比、前後比較 非実施仮説比 ランキング ベンチマーキング ポートフォリオ

出典：財団法人政策科学研究所作成資料

調査法には評価対象に関連する情報の質的側面に関する情報収集と、量的側面に関するデータ収集とがあります。情報収集として一般的なものが関係者に対するインタビュー調査や、評価対象の関連分野における専門家へのヒアリング調査、アンケート調査などがあります。また、同様の研究開発に関する国内外の事例調査、文献調査なども含まれます。評価対象の目的が社会的インパクトや経済効果にある場合には、顧客調査（市場調査）や社会調査が必要になるでしょう。

データ収集では、評価対象に関連する既往の公式統計や調査報告等に掲載されているデータにあたる必要があります。データが存在しない場合にはデータ化するための作業が必要になります。関係者に対するアンケート調査、マネジメント側が保持している申請・管理のデータ、文献情報からのテキストマイニング、データマイニングなどを実施し、望ましいデータを探索することになります。

データ収集に関して、特に対象の質的側面に対して何らかの測定基準を設け、その基準に照らし合わせて測定し、データ化する手法を「メジャメント (measurement)」と呼びます。

分析法の多くは、評価対象についての仮説検証を行うためのツールです。特に、既に測定されている統計データや数量データを加工し、評価対象の評価項目の内容を表すように分析結果を示したり、評価項目との関係性を裏付けるようなデータを得るような定量的分析手法のことを総称して「メトリクス (metrics)」といいます。メトリクスの多くは統計的手法を

用いるため、データサンプルが多くなければならないので、一般的には経年的、もしくはプログラムのように複数のプロジェクトに横断的な分析において使用されます。

表中のケース分析は事例比較から傾向を抽出する定性的な分析法です。コンテンツ分析はテキストマイニングの結果から統計的推論を用いて構成概念の枠組み作りを行います。計量文献分析（ビブリオメトリクス）は学術論文などを対象に成果の利用度指標を作成するのに用いられます。計量技術分析（テクノメトリクス）は特許データを対象に分析を行います。計量経済分析（エコノメトリクス）はコストデータや売上、社会費用などの経済データを分析の対象にする諸手法の総称です。計量社会分析（ソシオメトリクス）は社会的影響に関する指標化のツールです。これらの複合的な利用と発展的な分析方法に、各種多変量解析やシステム分析を組み合わせることになります。

メジャメントやメトリクスを活用するためには高度な専門知識を必要とするため、個別手法に精通したアナリストが分析の任を担います。

評価法にもさまざまな手法がありますが、これらのどれを採用するかという問題は、最終的な評価をどのようなかたちで行うか、という問題に置き換えることができます。一般的には、評価パネルを構成し、そこでさまざまな情報、データ、分析結果を勘案し、評価対象に総合的な評点をつけることになるでしょう。その意味では、表に挙げた定量的評価の手法群は分析法に含まれるかもしれません。

研究開発評価においてもっとも重要な評価法はレビュー法、パネル法です。科学技術の質的に深い側面を捉えるためには、ピアもしくはエキスパートの専門的知見による方法しかありません。ただし、その運営は非常にデリケートであり、高度なマネジメントが要求されます。

評点法（scoring）は本来定性的な評価内容を評定区分に照らし合わせて採点する方法です。ここで注意しなければならないことは、評点（score）それ自体は絶対的な意味を持つ数値ではないということです。これは、あるパネルの評点と別のパネルの評点を単純に足して2で割ることに意味がないことを考えれば自明でしょう。評点法の導入は、評価結果を明晰に示すことができるという点で望ましいものではありませんが、評点そのものが重要なのではなく、なぜそのような評点がついたのかという理由（コメント）を明らかにすることが重要であると理解して下さい。

定量的評価のうち、指標法は評価対象に関するさまざまな定量的表現を正規化し、指標（indicator）を作成して他の評価対象と比較する方法です。単一のデータから指標を作成する場合を単一指標、複数のデータを組み合わせて指標を作成する場合を複合指標といいます。

比較評価法は評価対象のパフォーマンスを相対的に評価する方法です。代替案がない場合には、評価対象の対前年比や前後比較、あるいは対計画年比をとります。プロジェクトやプログラムの関与が及ばなかったグループと比較して、評価対象のパフォーマンスを検証するコントロールグループ・アプローチや、プロジェクトやプログラムがなかった場合を想定して、それをベースラインとした時と比べてパフォーマンスがあったかどうかを実証的に検証する非実施仮説比があります。

ランキングはある指標に関して、評価対象群を上位から並べて比較する方法です。ベンチマーキングはひとこと言うと「それまでのベスト」と比較してどうかを見るという評価方法です。実験結果に関するチャンピオンデータなどのベストプラクティスを参照し、評価対象の優秀性を判断するのに用いられます。ポートフォリオは評価対象について、2つ以上の

指標を用いて評価座標軸を作成し、評価対象群の偏りを見るための手法です。

なお、こうした調査＝分析＝評価と、それらの結果を用いて行われる意思決定とは明確に区分しなければなりません。特に評価と意思決定を混同してしまうことで、被評価者を萎縮させてしまったり、負の評価が下されないよう防衛的にさせてしまい、評価が形骸化していくおそれがあります。評価の役割は、評価すべき対象が置かれている状況や対象の価値を明晰にすることであり、あくまでよりよい意思決定のための道具であることを忘れてはいけません。このことは何度強調してもよく、本テキストにおいても折に触れて言及します。

## 2) 定性的評価と定量的評価

先の表に示した調査法、分析法、評価法においては定性的な情報を扱う手法と定量的な情報を扱う手法とがあります。

評価に関わる誤解の一つに、「定量的評価＝客観的」、「定性的評価＝主観的」という考え方が根強くあります。定量的手法は一般的に結果の明晰性が高く、情報伝達の際の任意性が低いという面がありますが、反面、評価対象のある局面のみを限定的に切り取って数量化しているという特徴があります。

定性的手法は結果の表現を主に文章による叙述に頼ることになりますが、事象の本質を捉えることに関してはむしろ適切な場合の方が多いと言えます。

例えば、ワインの解説をする際に、ワインの甘辛度、酸味度、タンニン含有量、アルコール度等を指標化したところで、本当にそのワインが美味しいかどうかの判断材料にはなりません。むしろ、「ブラックベリーやプラム、カシスといった熟したフルーツの風味が・・・。」というような表現にこそ、飲む人の嗜好を判断させるための本質的な内容が含まれていると言えるでしょう。

通常、研究開発評価の俎上に上がるものは、

- 先行的事例に乏しく、比較不可能な事象を扱うケースが多い
- 一般化が困難で、マネジメントなどの経験的・学習的要素が多い
- 因果関係が明らかでなく、先行指標を得ることが難しい（妥当でない）

というような問題を含む対象であることがほとんどです。

特に研究開発評価では、成果の質に関わる評価をピアレビューやエキスパートレビューに委ねることになります。レビューは彼らの知識や経験を総動員してコメントを作成します。

したがって、「定性的評価だから主観的であり、客観性に欠ける（＝信頼性が低い）」ということにはなりません。

むしろ定性的評価において重要なことは、「どのような文脈で語られているか」ということです。歴史認識、比較対象、課題の喫緊性などが背景にあった上で、本質的な意味づけが行われることになります。

## (2) 定量的評価法の基本

### 1) 比較考量の基本的な考え方

ある研究開発のパフォーマンスを定量的に示すためには、比較の視点が絶対に必要になり

ます。代替案 (alternatives) が始めから存在していれば、比較は簡単でしょう。しかしながら、研究開発はプロセスの革新であれ、改良であれ、基本的に「これまでにない」モノを作り上げる行為に他なりません。そのため、比較考量をどこに求めるかが大変重要になります。

研究開発評価においてパフォーマンスを比較する場合の基本的な考え方を以下にまとめました。

### チャンピオンデータ

研究開発のアウトプットの性能等を比較する際に、もっとも直感的で分かりやすい方法は、当該分野のチャンピオンデータと比較するというものです。通常は、これまで世界一だったスコアを大きく更新したという場合にこの方法が用いられます。

図 3-21 は地球シミュレータの演算性能を比較した図です。地球シミュレータがそれまで世界最速であった米国のスーパーコンピュータ ASCI-White を実効性能で5倍も引き離れた事実をもっとも雄弁に物語っています。

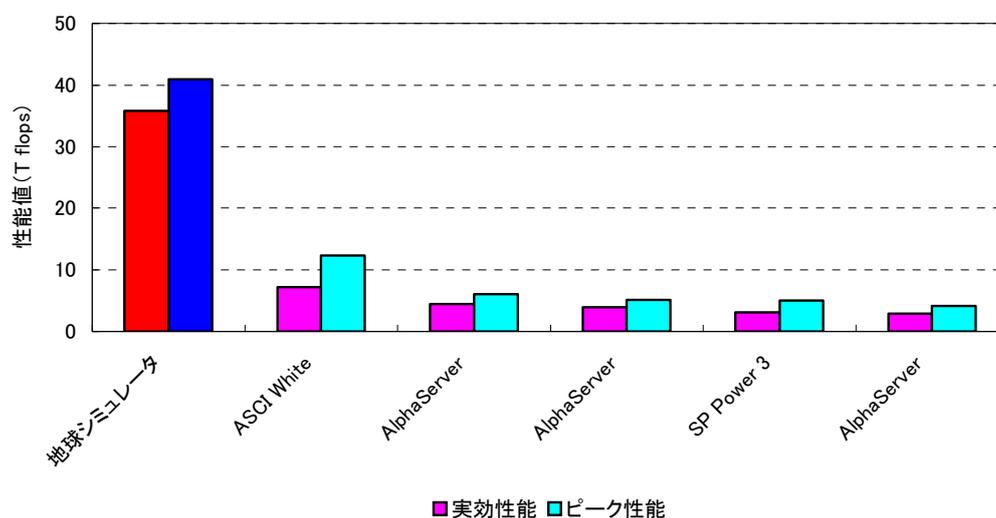


図 3-21 チャンピオンデータとの比較の例 (地球シミュレータの演算性能)

出典:財団法人政策科学研究所作成資料

また、世界一の更新でなくても、性能に関わる数値を費用対効果の指標として示すことで、コストパフォーマンスをチャンピオンデータと比較することもできます。他に比較可能な対象がある場合には、まずは性能指標等に関して直接比較を行うことが原則であると言えます。

### 前後比較 (“Before-After” comparison)

ある研究開発を実施する前と後とで、ターゲットにしている指標がどのように変化したかを見る比較の方法です。例えば、ある技術開発プログラムに参加した企業の製品技術が、プログラムに参加する前とした後で比較して、大幅に向上しているような場合に、それを当該

プログラムのパフォーマンスとして示す考え方です。

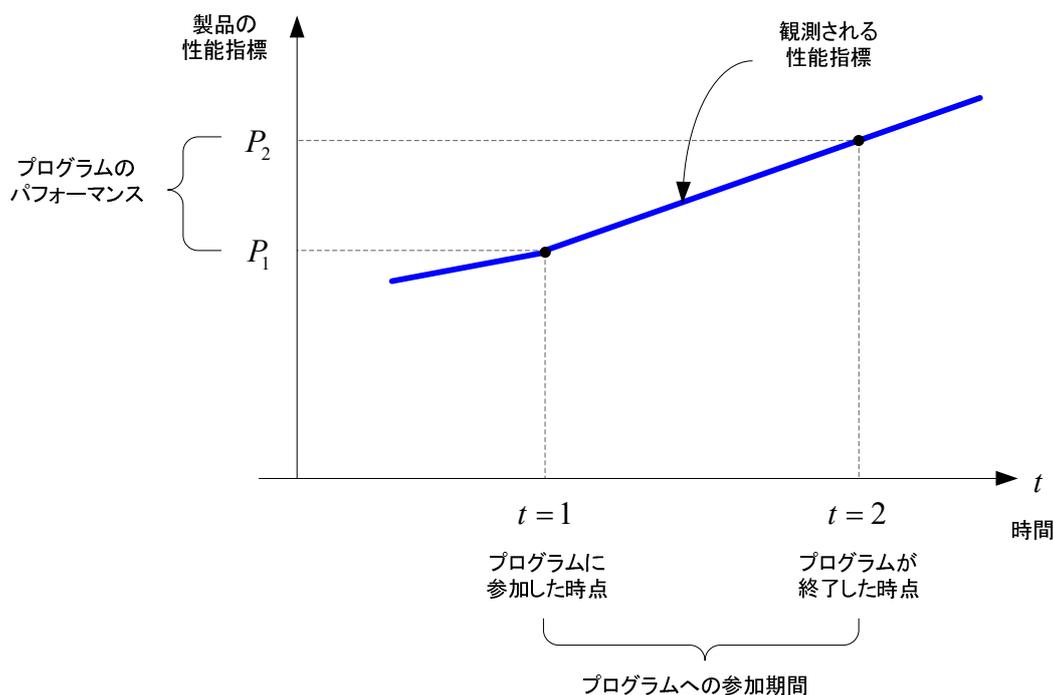


図 3-22 前後比較の考え方

出典:財団法人政策科学研究所作成資料

しかし、前後比較は当該技術分野に実績を有する企業がプログラムに参加している場合などには、プログラムのパフォーマンスとしての因果関係を保証する比較の仕方ではありません。その企業が元々行っていた研究開発の蓄積分や継続努力が結びついて性能を上昇させている部分と、当該プログラムによって研究開発が加速された部分とを切り分けて考えないと、本来のプログラムのパフォーマンスを示したことにはならないからです。

### 非実施仮説("With-Without" comparison)

前後比較が観測される性能指標のある一定期間内の向上分を示しているのに対して、非実施仮説による比較考量は、当該プログラムの「真の貢献」部分を示します。

先ほどと同じ例で考えてみましょう。技術開発プログラムに参加している企業は、当該技術分野における研究開発の蓄積があり、ある時期から国が実施しているプログラムによって研究開発資金を助成してもらい、研究開発を実施したとします。

ここで、当該技術開発プログラムの寄与度はどのように考えたらよいのでしょうか。もし、その企業が技術開発プログラムに参加せず、自前で研究開発を継続したとしたら、製品の性能指標はどのように推移していたのでしょうか。これを推計するための基本的な考え方が「非実施仮説」です。図 3-23 を用いて非実施仮説によるパフォーマンスの計測方法を解説します。

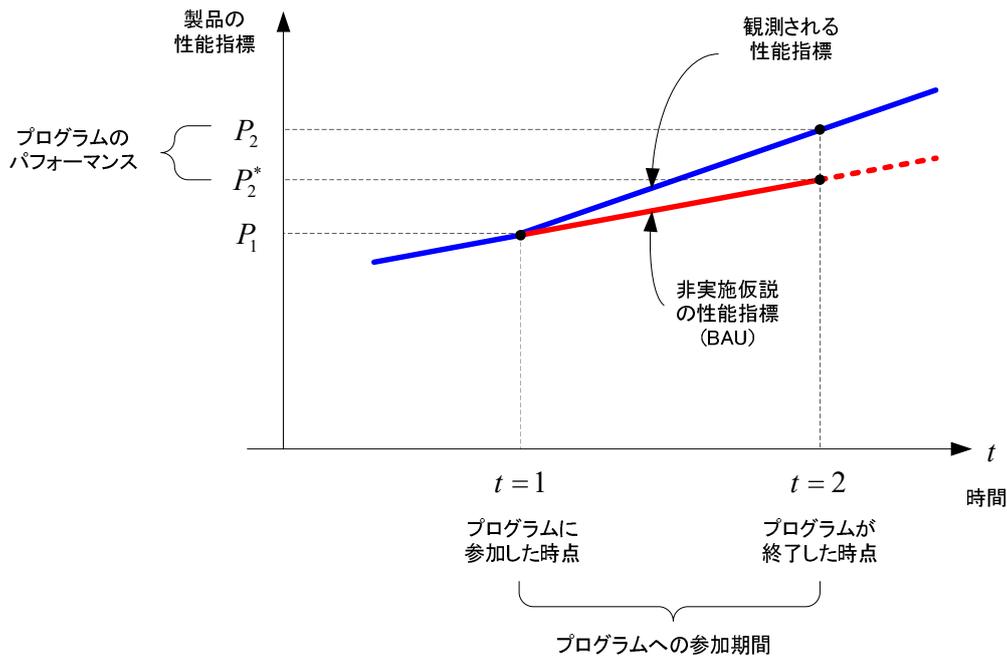


図 3-23 非実施仮説の考え方

出典:財団法人政策科学研究所作成資料

仮にそれまでの性能指標のトレンドがデータから得られたとしましょう。この時、企業が自前で研究開発した場合の達成したであろう性能指標の値は  $P_2^*$  になります。一方、観測された性能指標の値は  $P_2$  です。 $P_2^*$  と  $P_2$  の差は、技術開発プログラムに参加したことで、当該企業の研究開発が加速された部分と解釈できます。つまり、この差を技術開発プログラムのパフォーマンスとして示すことができます。

非実施仮説による比較考量は、継続性のある事象の評価では不可欠です。「もし、研究開発プログラムに参加していなかったら (Without)」という状況を、BAU (Business As Usual) ケースと呼びます。BAU で推計される指標と、「プログラムに参加した場合 (With)」の現実の指標との差が、非実施仮説における比較考量となります。

### コントロールグループ・アプローチ

非実施仮説では、文字通り実施しなかった場合を想定することで比較考量としていますが、実施しない場合を現実のサンプルに求めることもできます。この比較考量の求め方をコントロールグループ・アプローチと呼びます。このもっとも代表的な例が、新薬の治験（臨床試験）で行われる「比較試験」です。

新薬の治験では、プラセボ（偽薬）を用いて薬理効果を検証します。新薬を投与する患者のグループと、全く治療効果が無い「プラセボ」を投与する患者のグループ（プラセボ群）に分けて、それ以外は全く同じ条件で治験を実施し、それによる臨床データを測定し、両者のデータを比較して治療効果を確認します。プラセボは新薬と見た目も味までも同じに製剤されているので、患者はもちろんのこと、投与する側の医師ですら区別することができません。偽薬の投与によっても症状が改善してしまうというプラシーボ効果すらもデータのバイアスとして考慮した非常に厳密な比較試験です。

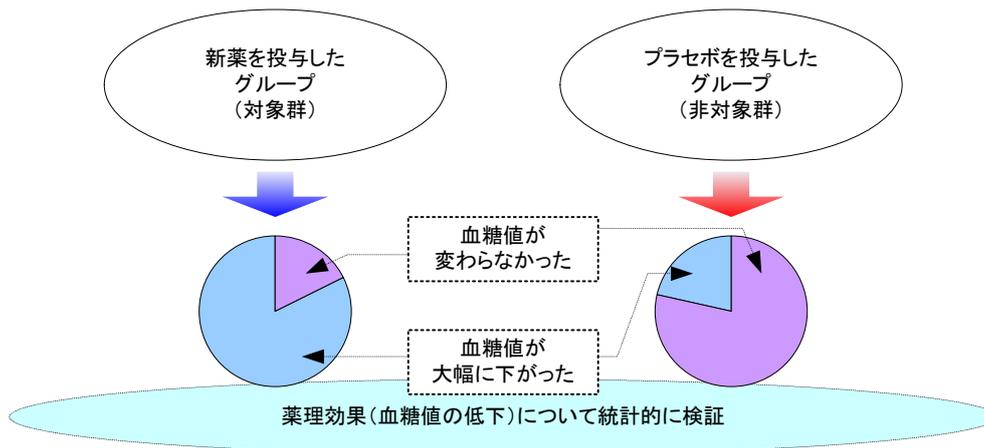


図 3-24 コントロールグループ・アプローチの考え方(例:新薬の比較試験)

出典:財団法人政策科学研究所作成資料

このとき、新薬を投与されたグループは対象群（コントロールグループ）と呼ばれ、プラセボを投与されたグループは非対象群と呼ばれます。対象群と非対象群との間に統計的な明確な差異が認められたとき、新薬の薬理効果が実証されるのです。

研究開発評価においてコントロールグループ・アプローチを使用する場面はサンプルを多く必要とするプログラムレベルでの評価に限定されると考えられます。例えば、プログラムに参加した企業と参加していない企業との間にどのようなパフォーマンスの差異があったかを追跡調査等で明らかにするような場合が相当します。

### (3) 数値データはどのように扱うべきか

私たちは普段気にすることなく、数値化されたものをデータと呼んでいます。しかし、データには種類があり、それらを理解しなければ誤った活用をしてしまうことになります。まずはデータの種類を学びましょう。

データには大きく分けて質的データと量的データとがあります。質的データは数値であっても四則演算が適用できないデータのことです。それに対して量的データは四則演算が可能なものといえます。

質的データはさらに名義尺度（nominal scale）と順序尺度（ordinal scale）に分かれます。名義尺度とは、カテゴリーの分類に使用されるデータのことです。その数値の順序、大きさには全く意味のない情報のことです。データベースのコード番号なども名義尺度の一つです。順序尺度とは大小関係や順位の相対的な比較に用いられる判定条件に対応したデータです。研究開発評価における評点は順序尺度になります。順序尺度は数値の値そのものに意味があるわけではないので、平均値を取ることができません。しかしながら、最頻値（mode）や中央値（median）を取ることで代表性を表現することが可能です。

量的データも間隔尺度（interval scale）と比例尺度（ratio scale）に分かれます。間隔尺度は距離尺度とも呼ばれますが、数値の差にのみ意味があるデータです。もっとも一般的なのは例えば0～1の範囲で正規化された指標です。つまり、パーセンテージ（%）で示されるデータです。間隔尺度では数値の相対的な距離を定めているだけですので、比を取ることができません。「+4ポイント」という表現はできても「3%上昇」という表現はできません。

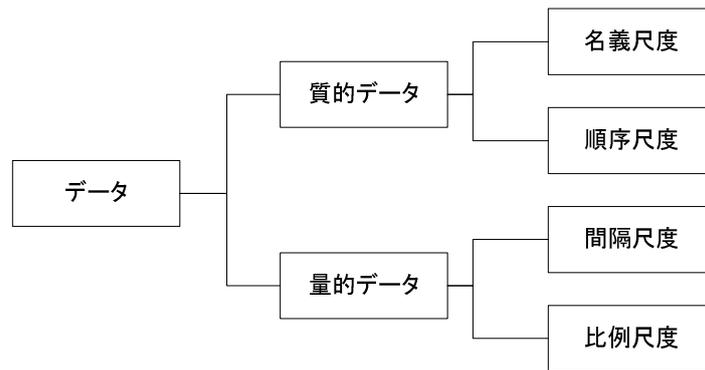


図 3-25 データの種類

出典：財団法人政策科学研究所作成資料

量的データのうち、比例尺度になってはじめて、四則演算の全てが適用できます。比例尺度のデータとは絶対的な値を持つデータのことです。つまり、比が定義できる絶対零点を持つデータです。

表 3-7 データの4分類

分類	説明	例	対応するデータ
名義尺度 Nominal scale	カテゴリーの分類に使用される。平均値を求めても意味がない。	血液型: A→1, B→2, O→3, AB→4 性別: 男→0, 女→1	質的データ
順序尺度 Ordinal scale	大小関係や順位の相対的な比較に使用される。数値の絶対的な大小に意味がないので、平均値を取ることはできない。中央値であれば可。	判定: とても良い→5, 良い→4, 普通→3, 悪い→2, とても悪い→1	質的データ ※ 順序や大小関係に等間隔を仮定すれば、間隔尺度として計算できる。
間隔尺度 Interval scale	距離尺度とも呼ぶ。数値の差にのみ意味があるだけなので、「%上昇」といった表現はできない。	温度(摂氏)、テストの得点	量的データ
比例尺度 Ratio scale	比尺度とも呼ぶ。数値の差とともに数値の比にも意味がある。比が定義できることは絶対零点を持つことを意味する。	温度(華氏)、距離、時間、重量	量的データ

出典：財団法人政策科学研究所作成資料

#### (4) 評価項目、評価基準、評定区分とその構造

評価作業を準備するうえで最も困難なことは、「評価項目」と「評価基準」そして「評定区分」とその構造を具体的に定めることです。これは評価体系そのものでもあり、評価の理念や枠組みをはじめ評価目的や評価法等を具体的な手続のかたちとしてまとめたものに相当します。現在我が国で実施されている公的資金による研究評価制度で、上記の三者が完備している制度はほとんどありません。特に評価基準に相当する内容が示されていないことが多く、「評価基準」の用語が用いられていても内容的には「評定区分」であったり、「評価項目」と「評価基準」を混同していたりします。多くの場合では、内容的にみて評価項目と評定区分の構造が示されているにすぎません。このような評価体系に従うとすれば、評価者は暗黙的に自己の評価基準を想定し評定を実施することになります。これでは評価結果が開示されたにしても、評価プロセスの論理が明示されていないので、暗黙的評価であると言わざるを得ませ

ん。

評価項目とは評価内容を区分しその区分カテゴリー毎に設定された評価局面を特定するための項目のことです。評価内容は、評価目的をブレイクダウンして得られます。具体的には、評価目的に照らして評価対象の状況を把握するために必要とされる事項を特定する作業に相当します。評価項目は互いに **MECE** な関係で構成されているべきものです。即ち項目間に重複が無くかつ項目全てを合わせると全評価局面を構成していなくてはなりません。

評価基準とは評定の判断を導く根拠のことであり、判断のための論理や指標あるいはその導出法等として設定されます。評価項目によっては評価基準が単一で自明のこともあります。通常は複数の基準が想定されるので、いずれの判断根拠を用いるのか、また複数の基準を採用するのか等を特定しておく必要があります。例えば、基礎科学の研究領域で、「研究の質」に関わる評価項目に対する基準として「独創性」とか「創造性」が示されていることが多いですが、このような概念では基準としての機能を期待できません。「独創性」や「創造性」の尺度があいまいであるからです。推奨されている評価基準は「知的フロンティアへの挑戦」や「知的フロンティアとしてのインパクトの大きさ」等になります。

評定区分とは評定結果を表示する際に用いる区分のことであり、通常階層構造や平面分割構造をしています。評価結果を文章で表現するだけの場合では直接的には評定区分は必要ありませんが、その場合でも評価結果を表現するための論理構造を示すことは暗黙的であるにしても必要です。評定区分はその論理構造を顕在化させたものとも言えます。評定区分はその内容を単に記号で表すだけでなく、その記号の意味内容を特定して表示する必要があります。「1」、「2」、「3」等の数字で表現された階層区分であっても、区分構造としては必ずしも等価に区分されているとは限りません。

さて、データの種類について理解が深まったところで、研究開発評価における数値データ活用の例を見てみましょう。

## **評点の場合**

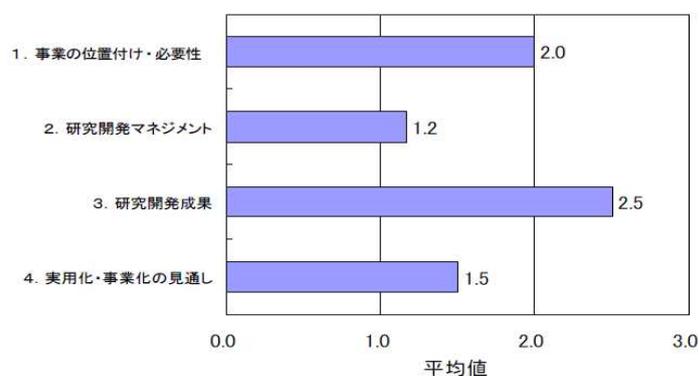
研究開発評価の評点について考えてみましょう。ここでは、新エネルギー・産業技術総合開発機構 (NEDO) における評点法の適用例を見てみましょう。古いデータになりますが、この場合の評点と評定区分の定義は以下のようになっています。

表 3-8 新エネルギー・産業技術総合開発機構の評定区分

判定基準		
1. 事業の位置づけ・必要性について		
・非常に重要 →A	考慮事項(参考) (1) NEDO、産総研の事業としての妥当性 (2) 事業目的の妥当性	[ a b c d ] [ a b c d ]
・重要 →B		
・概ね妥当 →C		
・妥当性がない、又は失われた →D		
2. 研究開発マネジメントについて		
・非常によい →A	考慮事項(参考) (1) 研究開発目標の妥当性 (2) 研究開発計画の妥当性 (3) 研究開発実施者の事業体制の妥当性 (4) 情勢変化への対応等	[ a b c d ] [ a b c d ] [ a b c d ] [ a b c d ]
・よい →B		
・概ね適切 →C		
・適切とはいえない →D		
3. 研究開発成果について		
・非常によい →A	考慮事項(参考) (1) 目標の達成度 (2) 成果の意義 (3) 特許の取得 (4) 論文発表・成果の普及 (5) 標準化(産総研担当分)	[ a b c d ] [ a b c d ]
・よい →B		
・概ね妥当 →C		
・妥当とはいえない →D		
4. 実用化・事業化の見通しについて		
・明確に実現可能なプランあり →A	考慮事項(参考) (1) 成果の実用可能性 (2) 波及効果 (3) 事業化までのシナリオ	[ a b c d ] [ a b c d ] [ a b c d ]
・実現可能なプランあり →B		
・概ね実現可能なプランあり →C		
・見通しが不明 →D		

出典：新エネルギー・産業技術総合開発機構作成資料

上の評定区分に従って、評価委員会の各委員が評点を付けます。その結果が下の図で開示されています。



評価項目	平均値	素点 (注)							
		B	B	C	B	C	C	B	
1. 事業の位置付け・必要性について	2.0	B	B	C	B	C	C	B	
2. 研究開発マネジメントについて	1.2	B	B	A	B	A	A	D	
3. 研究開発成果について	2.5	C	B	C	B	D	D	C	
4. 実用化・事業化の見通しについて	1.5	D	B	A	C	B	A	B	

(注) A=3, B=2, C=1, D=0 として事務局が数値に換算し、平均値を算出。

図 3-26 新エネルギー・産業技術総合開発機構の評定結果の例

出典：新エネルギー・産業技術総合開発機構作成資料

このような結果を得るためには、前提条件として、順序尺度である評点に対して間隔尺度

を適用しているという点が指摘できます。つまり、評点 A と B の差は C と D の差に等しいという解釈を前提としています。また、個人の評点の差についても、個人間で互いに等しいという解釈を前提としています。

評価結果を明示的に表すにはこのような表現方法もやむを得ないと思われます。しかしながら、より重要なことは、このような評価結果を得るに至る評価委員会でのマネジメントの在り方となります。本来、個人の順序尺度である評点を、評価委員会共通の間隔尺度として用いるためには、評点を導くための評定区分に対する理解が委員会でも共有される必要があります。評定区分に対して解釈の個人差が大きく生じないように、評価シートが設計されていなければなりません。

また、こうした評点を示すとしても、各委員の評点（素点）がどのような分布になっているかを補助的に示すことが大切になります。平均値だけを見た場合、分布が平均値の周辺に集まっている場合と、極端に評点が割れている場合との違いが判別できないからです。

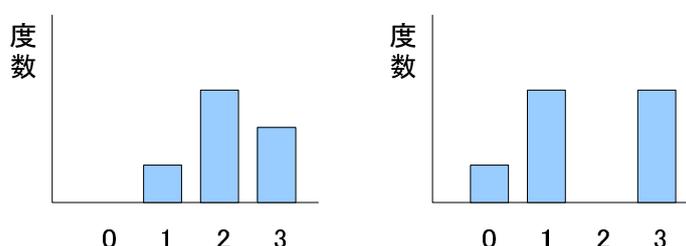


図 3-27 平均値が同じでも分布が異なるパターンの例

出典：財団法人政策科学研究所作成資料

表 3-9 経済産業省の技術評価結果(3点満点)の例

評点結果		
評価項目	平均点	標準偏差
1. 事業の目的・政策的位置付け	2.83	0.41
2. 研究開発目標の妥当性	2.33	0.52
3. 成果、目標の達成度の妥当性	1.83	0.41
4. 波及効果、事業化についての妥当性	2.00	0.63
5. 研究開発マネジメント・体制・資金・費用対効果等の妥当性	1.83	0.41
6. 総合評価	2.50	0.55

出典：経済産業省平成 18 年度技術評価報告書より抜粋

したがって、NEDO の例のように、各評価委員の素点を示すか、経済産業省の実施している研究開発プロジェクトの技術評価のように、平均点の他に標準偏差（値が大きいほど評点が割れている目安になります）を示すなどの工夫が必要となります。

参考までに、評定区分の例をいくつか示します（表 3-10）。

表 3-10 評定区分の例

	評定区分の内容	特色
2段階 (1, 0)	<ul style="list-style-type: none"> <li>合、否</li> <li>適切、不適切</li> </ul>	採否等を明確に表現するための総合評価等に用いる。
3段階 (A, B, C) (3, 2, 1)	<ul style="list-style-type: none"> <li>極めて良好、概ね良好、良好でない</li> <li>極めて適切、概ね適切、適切でない</li> </ul>	中間評価や直後評価の対計画比等の表現に用いる。ほぼ計画通りであるなら「概ね良好」、「概ね適切」とする。
(A, B, C) (2, 1, 0)	<ul style="list-style-type: none"> <li>非常に期待される、期待できる、期待できない</li> </ul>	事前評価に用いる場合、合格側を「非常に期待される」と「期待できる」に2区分し、不合格側を「期待できない」とする。従って合格ラインを50点とすると100-75, 74-50, 49-0の3区分となる。
4段階 (A, B, C, D) (4, 3, 2, 1)	<ul style="list-style-type: none"> <li>非常に優れている、優れている、劣る、非常に劣る</li> </ul>	「平均段階」をあえて設定せず、優劣を積極的につける場合に用いる。
	<ul style="list-style-type: none"> <li>極めて優れている、かなり優れている、優れている、劣る</li> </ul>	「優」「良」「可」「不可」のように、合格側(100-50)を3段階に区分し、不合格側(49以下)は1段階で表現する。
(AA, A, B, C)	<ul style="list-style-type: none"> <li>極めて良好、概ね良好、良好、良好でない</li> </ul>	3段階評価に特記部分(AA)を加える。「優れている」側のグレードに飛び抜けた開きがある場合。
5段階 (A, B, C, D, E) (5, 4, 3, 2, 1)	<ul style="list-style-type: none"> <li>非常に優れている、優れている、良好、劣る、非常に劣る</li> <li>HIGH、GOOD、AVERAGE、FAIR、LOW</li> </ul>	中間値付近を「良好」として等価に区分する。
(AA, A, B, C, D)	<ul style="list-style-type: none"> <li>極めて優れている、かなり優れている、優れている、劣る、かなり劣る</li> </ul>	4段階評価+特記部分(AA)の例。
6段階評価以上 (10段階評価の例)	(10, 9, 8, 7, 6, 5, 4, 3, 2, 1)	上記をさらに細分化する。 例えば10段階では、10-8.5(HIGH)、8.4-6.5(GOOD)、6.4-4.5(AVERAGE)、4.4-2.5(FAIR)、2.4-1(LOW)等の区分に相当するものとする。

出典：研究開発プロジェクト等の評価手法に関する調査報告書(平成13年度経済産業省委託調査)

### 費用便益比の場合

研究開発評価において、費用便益比を計算することは極めて難しいと言えます。事業終了直後では研究開発投資の効果が出現していることはほとんど希だからです。また、さまざまな効果を金額換算する方法も用意されていますが、一律に適用できるような換算値は存在しません。しかし、何らかの予測を立て、かつ高度な金額換算手法を用いて便益が計算できたとすれば、研究開発投資費用に対する便益比が計算できます。

問題は、このような比率の持つ数値的意味です。費用便益比は当該プロジェクトの経済性効率(パフォーマンス)を見ているだけであって、利益の大きさを見るものではありません。また、分子と分母がそれぞれ独立に計算されているものを対比した複合指標であるため、他のプロジェクトとの比較において絶対基準を持ちません。したがって、非常に高度な計算を行い、定量的分析を厳密に行っているように見えても、比較の観点からは、費用便益比は間隔尺度と同じと言えます。

つまり、プロジェクトが経済効率的(この言葉の意味も本来は深く考える必要があります)に実施されたかどうかについて、他のプロジェクトとの相対的な比較の観点なしに「大

変効率的に実施された」、「ある程度効率的に実施された」、「効率的に実施された」とは言い難い」というような評定区分を客観的に数値化したということと同じ意味合いしか持ちません。

費用便益比を比較可能な数値として意味づけるためには、「投資の機会費用」という考え方が不可欠です。つまり、投資の一部を金融機関からの借り入れで賄う場合、かつ、事業に財務的収入が事業の想定している計画年限に見込める場合においてのみ意味を持ちます。なぜならば、この場合、費用便益比（あるいは内部収益率）が大きいほど、投資の回収期間が短い、つまり収益を生みやすいという説明になるからです。ですから、民間の研究開発投資のために試算することには一定の意味があります。

翻って、国が行う研究開発投資においてすべからず費用便益比を計算する意味がどこまであるのでしょうか。個別の研究開発プロジェクトの目的にもよりますが、おそらくは計算する労力が大きいだけで、実務的な意味合いからは全くないと言っても過言ではないでしょう。あるとすれば、事業管理主体の財務当局へのエクスキューズか、財務当局自身の国民へのエクスキューズという意味合いしか残らないでしょう。そうした数値が保証するものは、「お金を無駄に使ったわけではない」という意味であって、「どれほど役に立ったか」を明らかにしていないからです。

心ある経済学者は費用便益分析よりも費用効果分析を推奨しています。費用効果分析では、効果に相当する数値を物理量であらわすため、単純に何に役に立ったかを明らかにしやすいからです。例えば、温暖化対策のための研究開発投資であれば、成果がもたらす CO<sub>2</sub> 削減量 1 トンに対する投資額が明確になります。3,000 円/t-CO<sub>2</sub> という数値であれば、他の研究開発プロジェクトとその効果を比較することができるからです。この場合の数値は比例尺度である言えます。

## (5) 研究開発の「質」に関わる評価法

一般的に、研究開発の「質」を評価するためのもっとも有効な方法はピア、エキスパートによるレビューです。我が国の研究開発評価の制度は、早くから外部の専門家、有識者等による評価を導入し、評価委員会や評価分科会などの評価パネルが研究開発評価の主軸を担ってきました。

しかしながら、レビュー評価法は研究開発の、とりわけ成果の「質」については有効な評価方法ですが、その他の評価項目に関して万能に機能するわけではありません。また、その「質」に関しても、ある単独のディシプリン内の評価で済むものから、学際的な評価が必要なもの、社会経済的影響が大きいものなどさまざまです。

つまり、学識経験者や有識者と呼ばれる方々を単純に配置すれば良いのではなく、評価対象の特性に応じてレビューの仕方やレビューアの構成を考えなければなりません。また、評価パネルを設置し、委員長を任命したら自動的に審議と評価が行われるのではなく、パネルの運営も重要な仕事になります。

ここでは、原則論としての評価パネルの在り方を考えてみましょう。

### 1) 評価対象に応じたレビューの仕組み

評価対象の目的に応じて、構成されるレビューアのタイプが異なります。

一般に、レビューアのタイプにはピア (Peer) とエキスパート (Expert) があります。

ピアの原義は、「仲間内、同業者」という意味です。研究開発評価におけるピアレビューの意味は、あるディシプリン（学問の一分野）内における専門的知識を持った研究者仲間内での評価ということになります。

一方、エキスパートは「有識者」と言えば良いでしょうか。単一のディシプリンだけでなく、複数の学問分野、あるいは技術領域、法制度、経営、マーケティング等に関する幅広い知識を持ち、研究開発の成果に対して複眼的なものの見方ができる人材がこれに相当します。これを担うのは研究者だけでなく、実業家、専門家、行政職等の人材であり、これらの人々がプログラム等の目的に応じて配置されます。

表 3-11 ピアとエキスパートの適用区分

	科学技術的価値 の側面を重視	社会経済的価値 の側面を重視
単一ディシプリン	・ピアレビュー ・ピアパネル	
複合的ディシプリン	・エキスパートレビュー ・エキスパートパネル	・エキスパートレビュー ・エキスパートパネル ・エキスパートボード

※ディシプリン(Discipline)の原義は「しつけ」。学問領域毎に専門的知識を見つける「しつけ」や「訓練方法」が異なるため、それらが機能する「特定の研究領域」の意味に用いられるようになった。

出典：財団法人政策科学研究所作成資料

上の表 3-11 はピアとエキスパートの適用区分についてまとめたものです。まず、科学技術的価値を重視した評価を行うことを前提に考えてみましょう。単一ディシプリンの場合であれば、当該分野の研究者内でのレビューをもとに評価を行うしかありません。一方、学際的領域や新領域など複合的ディシプリンの場合、さまざまな領域の専門家が集い、お互いの評価について話し合うといった工夫が重要になります。その際、評価者は各々評価を行うにせよ、パネルリーダーはそれらの言い分に理解のある、幅広い見識を持った人材（エキスパート）でなければなりません。

また、社会経済的価値の側面を重視した評価を行う時には、単一ディシプリンのピアレビューは原則として意味を持ちません。社会経済的価値そのものが、さまざまな学問分野の関心領域でもあるからです。むしろ、ここで問題となるのは、より評価の難しい、社会的インパクトについての判断など、専門家の判断だけによらない評価の枠組みが新たに必要となる点です。表中、エキスパートボードとあるのは、評価に最終的な責任を負うことの出来る、意思決定者により近い存在の有識者の集まりです。

詳細については後述しますが、特にエキスパートによるレビューでは、パネル構成やレビューアの選出方法がマネジメントの重要な点になります。

## 2) パネル運営

レビューアからなる委員会(パネル)を組織し、評価を行う方法をパネル評価と言います。

評価パネル方式では、異なる専門家の意見を聞くことが前提ですので、もっぱら学際的研究か、新領域の研究、もしくはミッション型の研究開発が評価対象となります。つまり、評価パネル方式では少なくともエキスパートが参画することが前提となります。

ここでは評価パネルの運営の原則を学びましょう。以下の図 3-28 の流れでは、次のよう

なアクターが登場します。

評価委員会は、評価パネル（パネルリーダーを含む）と評価事務局から構成されます。場合によっては、被評価者（研究開発実施者や事業推進者）がこれに参加することもあります。

評価パネルは評価事務局から示される評価項目・評価基準に従って評価を行います。そのために必要な情報やデータ、場合によってはヒアリングや現地見学等から得られた情報を下に評価を行います。評価事務局は評価パネルの選定、委員会招集のロジスティクス、評価項目・評価基準の作成を行います。また、被評価者に対しては評価方針の事前説明や、評価結果の通知を行います。被評価者は、中間・事後評価であれば、評価に先立ち、事業等の成果報告書を提出したり、さらに評価が公正に行われるように評価委員会に対して評価に必要なデータの提供を行ったりします。また、評価結果に異議がある場合、申し立てを行うこともあります。

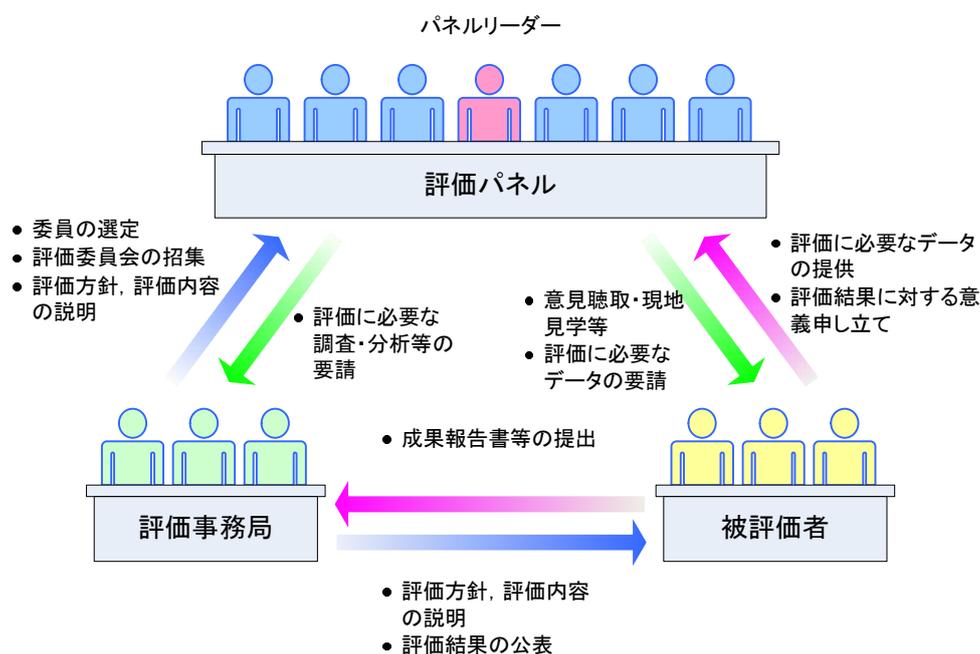


図 3-28 評価委員会のアクター

出典: 財団法人政策科学研究所作成資料

### レビューアの選定

評価パネルのレビューアを選定します。評価対象の科学技術領域や社会経済領域にしたがって、学識経験者、民間の技術者、その他のステークホルダーを選定します。その他のステークホルダーとしては、より広い見知から意見のできる科学ジャーナリストや関連技術領域の企業人等があげられます。

この中で、専門領域を別にして、評価経験の有無からおおよそのレビューアの構成を考えます。例えば、7人のレビューアを選定するとしたら、5人程度は評価経験がある人材、2人程度は初めて評価を担当する人材というように配置します。これには、レビューア自身も評価経験を積むことで、より有効なレビューを行うことができるという理由と、レビューアのパイを増やして、研究開発評価のレビュー人材の厚みを増す必要があるからです。新人のレビューアは、過去の評価対象から優秀な評価を受けた研究者等や、他の有識者からの紹介

というかたちで発掘することができます。

### **パネルリーダーの選出**

評価事務局としては、選定したレビューアの中から、評価経験の豊富で、かつ、幅広い知見を有した人材（エキスパート）に対して予めパネルリーダーの任を担ってもらうよう、依頼をする必要があります。そして、評価委員会を複数回行う必要があるときには第1回目の委員会において、他のレビューア全員から承認を受けるかたちでパネルリーダーを正式に選出することが望ましいでしょう。

パネルリーダーはさまざまな意見に耳を傾け、議論の方向性が評価の観点とずれているようであればこれを修正するなどファシリテーターの役割を担います。また、最終的な評価がレビューアの中で割れてしまった場合には、パネルリーダーの権限で評点を決定するなど、他のレビューアに対して、評価パネルをコントロールする立場にもあります。したがって、民主的な手続を踏まえて選出されることが原則として必要です。

### **評価方針等の説明**

評価委員会においては、評価事務局からレビューアに対し、どのような評価を行って欲しいのかについてのレクチャーをする必要があります。ここで、評価事務局として評価方針の説明、評価対象に関する評価項目・評価基準を記した評価シートの説明を行います。評価シートに対する委員からの疑問、修正要望等があれば、この段階で修正し、評価委員会において共通の理解を得る必要があります。

### **評価対象の説明と情報・データの確認**

プログラム評価や比較的規模の大きいプロジェクトの評価の場合、評価委員会において、評価対象の説明を被評価者が行うことがあります。小額のプロジェクトに係る事前評価（採択審査）の場合には被評価者は委員会に同席せず、提出された応募書類のみに基づいて評価を行うことが通常です。

評価パネルは評価項目や評価基準に照らして、判断に必要なデータに不足があれば、被評価者にデータや情報の提供を求めます。また、アウトカムに関する指標や分析結果が必要であれば、評価事務局に対して調査・分析の要請を行います。

また、評価対象となるプロジェクト等の規模や性格によっては、オンサイトの評価を実施する必要があります。拠点形成や実験設備、大型の研究開発基盤に関わる評価であれば、現地の見学やヒアリングが不可欠でしょう。

なお、評価パネルは評価委員会で得た情報・データ等を他者に教えてはならないなどのルールが必要な場合があります。特に、特許などの知的財産に絡むような成果を扱う場合には、あらかじめ、そのようなルールを設定し、レビューアに誓約してもらうことが重要です。

### **審議**

評価パネルによる審議の段階では、被評価者は退席し、委員の間で意見交換をした後に評点付けなどの作業に移ります。書面審査によってあらかじめ各委員が評点付けしたものを持ち寄り、議論を行う場合もあります。パネルリーダーが主導的な役割を果たすのはこの段階

です。意見交換の後、やはり評点が付けられないようであれば、時間的な制約もあるため、パネルリーダーに一任するなどの措置が採られる場合もあります。

### 評価報告書の作成

評価事務局は評価結果である評点の集計と、コメントの集約作業を行います。通常、評価報告書案を作成した後、被評価者と評価パネル双方に対して報告書案を開示します。プログラム評価にせよプロジェクト評価にせよ、事前評価の段階ではあまりこうしたことは行われませんが、被評価者は評価報告書に記載された事実関係に間違いがある場合、その内容によって評価結果に影響をもたらしていると考えられる場合には、評価パネルに対して異議申し立ての機会を与えられて然るべきでしょう。評価パネルも評価報告書に記載されたコメントが誤って解釈されていないかどうかをチェックし、双方合意の上で承認され、評価報告書が完成します。

### 評価報告書の公表と意思決定

評価報告書案が承認された後、Web等で評価報告書として公開します。評価報告書には、評価パネルの氏名や肩書きが明記されているので、評価事務局のみならず、評価パネルも評価結果に対して責任を負う立場にあります。ただし、評価と意思決定は原則的に別のものです。評価結果は当然尊重されるべきものではありませんが、合理的な理由があれば、評価の結果を受け取った意思決定者が評価報告書に記載されている勧告等に従わない場合もありえます。評価の目的にもよりますが、評価結果を受けてどのような改善を誰の責任の下でいつまでに行うのか、また、評価結果とは異なる判断をした場合にはその理由は何かといった評価結果の活用に関する文書などもあわせて公開することで、評価の形骸化を防ぐことができるようになります。

#### コラム 5: グループダイナミクス問題

対面によるパネル評価の場合、相補的ないし創発的なプラス効果がある一方で、好ましくない効果も少なからず発生します。こうした問題は「グループダイナミクス」問題と呼ばれており、よく知られる現象としては、競って欠陥・誤謬・問題点探しに陥る「スタンドプレー効果」、少数意見を抑圧する「集団思考現象」、結論になりそうな意見に賛成する「寄らば大樹現象」、多数の中の一員であることによる「責任回避現象」等があげられます。その結果、正統主義やリスク回避の傾向が強くなり、真に革新的、挑戦的な提案が低く評価されることとなります。

このような欠陥に陥らないためには、レビューアの自覚や資質の向上に努めることも必要ですが、より基本的にはパネルリーダーのリーダーシップが重要です。パネルリーダーにはまさにファシリテーターとしての資質、能力が求められていると言えます。

米国の評価研究者である **Bozeman** らは、こうした問題への対処を含め、ピアレビューに基づく評価の改善のための指針として、次の 7 項目を提示しています。長くなりますが、翻訳したものを掲載します。

**指針 1：ピアレビューを他の評価法と組み合わせて用いること。**ピアレビューは主観的・解釈的・個人的なアプローチであり、専門家各人の持つ知識を最大限活用するが、同時に各人の持つ偏り、気質、死角を含むことも避けられない。ピアレビューの主観性は内在的なものであり、利益となる場合もあるのであって、それを減少させようとするよりは主観性の少ない他の指標で補完すべきである。この目的のためには客観的指標（論文数、引用数、特許件数などの生産性の尺度）の集合に基づくアプローチが最も適している。

**指針 2：ピアレビューを公開性の高い R&D 活動の評価に用いること。**研究の大部分が基礎研究に近い場合でも、短期的な目標や固有技術との関連が明らかな場合にはピアレビューによる評価は適切でないことが多い。このような場合にピアレビューを行おうとすれば評価者すなわち競争者であるから、最良の場合でも感情的対立の可能性があり、悪くすれば訴訟に至る可能性もある。

**指針 3：評価者の選定が容易であること。**評価者が得にくい理由には次のようなものがある。

- (1) R&D 評価の実施責任者が当該科学技術分野との交流が乏しい。
- (2) 評価される分野が小規模で特殊であるため適当な評価者を得ることが困難または不可能である（同分野の研究者が参加を拒むこともある）。
- (3) 同分野の研究者が分散しており相互交流が少ない。産業界では特にこのようなことが多い。
- (4) 分野が新しいため信用を保証するメカニズム（専門誌、学界など）が整備されていない。評価者の選定の難易を問わず、評価者の適切性に何らかの疑問があればピアレビューによる評価自体が疑わしいものになる。したがって適切な評価者が得られる見込みがないか、参加に対する積極性が認められないならば、他の評価方法を採用すべきである。

**指針 4：内部の評価者を避けること。**評価をインサイダーとアウトサイダーのどちらが行うべきかは評価や政策に関する文献で多くの論争を呼んだ主題であり、どちらにも明らかな利点が存在する。しかし R&D の評価においては、外部の評価者を起用する利益は他の場合にまさって大きい。組織において、特に産業界の大組織において、自らに属する研究者に他部門の研究者の業績を評価させることはかなり普通に行われているが、この方式は正規の評価方法としては推奨できない。このような内部のピアレビューにもそれなりの用途はあるが、評価結果が意思決定に影響するような場合は、内部ピアレビューは同系交配のようなものであり、視野が限られることになる。アウトサイダーを起用することは専門性を強化することに役立つだけでなく（組織内に十分な専門家が存在する場合もある）、視野や職業的・組織的文化の多様性を強める効果がある。内部ピアレビューは、廃止する必要はないが、正規の R&D 評価手段として用いてはならない。

**指針 5：脱機能的グループダイナミクスを防止すること。**ピアレビューによる R&D 評価におけるグループダイナミクスの研究例は僅かであるが（たとえば Porter & Rossini, 1985）、他の場合に極めて普通に見られる問題点をピアレビューのパネルのみが免れていると考える理由はない。グループ思考、「寄らば大樹」現象、声の大きい者が勝つ現象などのグループダイナミクスを防止する可能性のある方法として、ピアレビューをコンピュータ化し、匿名の評価者がリアルタイムで発言することが試みられているが（Bozeman & Bretschneider, 1990 など）、コミュニケーションの豊富さが減殺され、立ち上げが面倒であることが大きな

欠点である。むしろ単に、このような現象が起こる可能性を意識し、それを防ぐために常識を働かせる方がよいであろう。たとえばパネルディスカッションを始める前に各評価者が個人的な評価を予め行っておくことは有用であり、これら個人的評価が匿名でパネルディスカッション以前に配布されればなお良い。

**指針 6: 尺度を用いるときはその妥当性と信頼性を検証すること。** NSF や NIH の使用した評価尺度は極めて単純なもので、次元数も少なく、少数の段階を含む段階評価を行うにすぎない。このような単純さはある場合には望ましいものである。しかし政府の研究助成機関で自らの使用する尺度の妥当性や信頼性を検証しているところはほとんどない。ピアレビューで評価尺度を用いるときは、次のような問題のいくつかを検討しておくべきである。まず、評点は信頼できるか（同じ案件を繰り返し評価したとき、各評価者は前回と同じ評点を与えるか）。評価のずれは本質的に偶然誤差である。また、評価者間の信頼性はどの程度か。評価に大きな差があるならば、ピアレビューを利用するに十分なコンセンサスが達成されていない可能性がある。

定量的尺度は適切か。100 点満点の尺度を用いるのは、もし評価者がそのレベルの詳細さで質を識別できるのでなければ無意味である。反対に少数段階の評点では識別力が不十分な場合もある。尺度の適切さは経験的な問題であり、検証が必要である。評価の各次元の間にはどのような関係があるか。通常は各次元に等しい重みを与えられるか、または評価者にとっての重要性に応じた重みを与えられている。しかし次元間の多重共線性すなわち重複は統計的に検定すべき問題である。もしある次元での変動が本質的にすべて他次元で説明できるなら、その両次元に同じ重みを与えるのは適切でないであろう。

そもそも定量的尺度を使用する必要があるか。この問題には明確な解答は存在しない。まず必要なことは、評価者が定量的な評点をつけることを便利と考えるかどうか、またこれに関連して、用いられる評点システムを適切と考えるかどうかを知ることである。定量的評点の使用は評価の目的と関連していなければならない。多数のプロジェクトに関して助成継続の可否や助成のレベルに関連する決定を下すためには、定量的尺度は有用である。しかし単に進行中のプロジェクトの改善が目的であるならば、定量的評価はほとんど役に立たない。いずれにせよ、定量的評点には文章による説明をつけることが望ましい。これによって数値の意味や評価者の判断基準が明らかになり、また改善すべき点が示唆される可能性もある。

**指針 7: 評価者に偏りに関する申告書を出させること。** National Academy of Science のパネルでは、まずパネリストが個人的な偏りや利害衝突の可能性に関する広範囲にわたる申告書を提出する。利害衝突に関しては明確な規則が大部分の政府助成機関に存在する。このような申告は、偏りの危険をすべて避けるには至らないとしても、部分的には有効な手段である。

参考文献：

研究開発プロジェクト等の評価手法に関する調査報告書（平成 13 年度経済産業省委託調査）  
Barry Bozeman, B. & Melkers, J.(eds.), *Evaluating R&D Impacts : Methods and Practice*, Kluwer: Boston, 1993.

### 3.4.3 研究開発評価に関わる専門人材

ここでは研究開発評価に関わる専門人材（Who）について学びます。

研究開発評価は単に大学の先生を呼んできて、評価パネルを構成し、審議してもらっただけ

では成り立ちません。研究開発評価には、その役割に応じて専門的能力を発揮する高度な人材が必要となります。

研究開発の評価活動に関わる人材は、その人材が担うべき評価機能により大別すると、レビューア（reviewer）、プラクティショナー（practitioner）、アナリスト（analyst）の3種類に区分できます。

表 3-12 評価人材の分類

分類	レビューア reviewer (evaluator)	プラクティショナー practitioner	アナリスト analyst
定義	<p>評価パネルを構成し、評価対象の質的側面を専門的観点から明確にする。評価対象領域の専門的人材であり、エバリユエータとも呼ばれる。</p> <p>ディシプリン内部の評価に携わるピアレビューアと、学際的ないし実務的内容に関する評価に携わるエキスパートレビューアとがある。</p>	<p>行政関連機関内部で評価の実務や運営に携わり、評価運営の実務的専門性を有する人材。</p> <p>行政一般を担ういわゆる「ジェネラリスト」が、評価に係る組織内でのOJTや外部での教育・研修等の機会を経て評価の実務的専門性を獲得し、「プラクティショナー」と呼ぶに相応しい実務的評価人材に成長する。</p>	<p>評価対象を分析するための高度な手法を活かし、評価対象の実態を深く把握し、評価作業を専門的見地から遂行する人材。</p> <p>評価に係る「スペシャリスト」であり、深い評価活動や経験等の研鑽を経て、「プロフェッショナル」と呼ぶに相応しい高度な手法を駆使できるようになる。</p>

出典：財団法人政策科学研究所作成資料

ここでは、特に欧州で通例的に使用されている人材分類の考え方に従って、研究開発評価人材の役割とキャリアパスについて解説します。

### (1) レビューア

最初に、評価パネルを構成するレビューアについて解説します。

レビューアとは研究開発の質的側面に対して評価を行うことができる専門家です。多くは、研究開発の成果が属する専門領域における専門家であり、大学等の研究者、研究機関の研究者がその任を担います。

レビューアは採択評価の場合であれば、研究プロポーザルに目を通し、当該研究の新規性、独創性等を評価します。中間・事後評価では、産み出された成果に対して専門家としての評価を行います。

レビューアは組織を代表する者であってはなりません。あくまでも個人が主体です。同じ研究領域の発展、人材の育成のために社会的貢献活動の一端として、その役割を担う存在といえます。したがって、レビューとはボランティアベースで行われるものであるという認識がレビューアと評価実施側双方に必要です。

また、レビューア自体の育成・蓄積については、何かしらの方法論があるわけではない、ということも認識すべきでしょう。彼らは自分たちが属する専門領域でのプロフェッショナルであって、レビュー・システムとはそうしたプロの観点からみた評価を重視するという仕組みです。評価を行うための能力・見識はレビューアとなる前の段階、すなわち研究者としてのキャリアによって育まれるものです。評価実施側ができることは、そのようなプロに対して、どのようなレビューを行って欲しいかの具体的な「依頼」をすることです。

## (2) プラクティショナー

### 1) プラクティショナーとは

プラクティショナーとは行政関連機関（省庁、独立行政法人を含む）で研究開発評価の実務や運営に携わり、評価運営の実務的専門性を有する人材のことです。典型的な職種としては、プログラムの運営一般に携わる「プログラムオフィサー（PO）」がこれに該当します。

しかしながら、POのすべてがプラクティショナーというわけではありません。行政関連機関の内部職員は多くの場合、ゼネラリストであり、研究開発評価のための専門の人材と呼ばれるためには知識と経験を必要とします。

表 3-13 は米国の州立研究機関のプログラム・マネージャーの中途採用に関する応募要件です。プラクティショナーと呼ばれる人材の必要要件がどのようなものか参考に見てみましょう。

表 3-13 プラクティショナーの応募要件（評価マネージャーの中途採用例）

<p>職位 評価マネージャー(Evaluation Manager)</p> <p>内容 研究開発評価ディレクターの下、機関の 1 つ以上の研究開発プログラムの評価を担当します。内部スタッフや外部協力者、研究開発実施者とともに、評価計画やモニタリングシステムの開発・改善を行います。</p> <p>必須条件</p> <ul style="list-style-type: none"><li>・ 社会科学、公衆衛生もしくは関連する領域における修士の学位を持つ者</li><li>・ プログラム評価の経験が3年以上あること</li><li>・ 質的データか量的データもしくはその両者の評価・分析手法に関する知識を持つ者</li><li>・ 質的・量的データの分析経験を有する者</li></ul> <p>【基本的な業務内容(責任分野)】</p> <ol style="list-style-type: none"><li>1. プログラム評価全般のマネジメント<ol style="list-style-type: none"><li>a ロジックモデルや評価計画の開発</li><li>b データ収集ツールの開発</li><li>c データ収集作業の調整</li><li>d データ入力やデータ統合の管理</li><li>e 量的・質的データの分析</li><li>f プログラムのモニタリングシステムの設計</li><li>g 評価結果のフィードバックメカニズムやレポートシステムの開発</li><li>h 評価手続や評価方法論の一貫した適用</li></ol></li><li>2. 内部スタッフや外部協力者と協働した評価能力形成の支援<ol style="list-style-type: none"><li>a 評価戦略やデータ収集、パフォーマンス測定に関する内部スタッフの指導</li><li>b 研究開発実施者に対する評価の技術的支援の提供</li><li>c 評価やデータ活用における能力開発の管理</li><li>d 内部スタッフに対する政策決定や意思決定における評価データ活用のための指導</li><li>e 外部協力者との効果的な共同作業</li></ol></li><li>3. 評価スタッフと研究開発実施者のマネジメント<ol style="list-style-type: none"><li>a 評価スタッフの統括</li><li>b 外部発注のマネジメント</li><li>c 研究開発実施者との契約とモニタリング</li></ol></li><li>4. 監査と成果の普及開発<ol style="list-style-type: none"><li>a 監査報告書の作成</li><li>b 内部スタッフとの協働による出版物等の普及啓発の促進</li><li>c 他の評価スタッフによる報告書の統括</li></ol></li><li>5. 評価委員メンバーとしての役割<ol style="list-style-type: none"><li>a 評価委員会への貢献</li><li>b 機関ミッションに関するワークグループや戦略企画会議への参加</li><li>c 評価計画及びマネジメントに関する評価委員会との必要に応じた協議</li><li>d プロポーザル開発や資金獲得業務に対する支援</li></ol></li></ol>
---

## 6. 望ましいスキル、知識、態度

- a 口頭並びに文書によるコミュニケーション・スキル
- b チームワークと良好な対人関係
- c プロジェクト・マネジメントのスキルとグループの調整スキル
- d 細心の注意力
- e 前向きな思考と自主的な動き
- f 多種多様な業務をこなすことのできる体系的スキル
- g サーベイやデータ収集ツールの開発経験
- h SPSS や SAS などの主要な統計ソフトウェアを用いた分析経験
- i Office ソフトを十分に使いこなせるスキル

### 給与(年俵)

マネージャー職 5,000~6,000ドル

シニア・マネージャー職 6,000~7,500ドル

出典: <<http://www.monster.com/>>より財団法人政策科学研究所作成

この州立研究機関では公衆衛生を改善するための研究開発プログラムを運営しており、プログラム・マネージャーに必須の要件として、公衆衛生に関する修士以上の学位を要求しています。一方で、公衆衛生のような社会的インパクトを評価するためには、社会科学の知識が不可欠です。したがって、社会科学の修士以上の学位が第一に要求されています。

また、データ分析に関する知識と経験も要求されています。この州立研究所では、内部スタッフ以外に外部の協力者と連携して評価業務を遂行していますが、データ分析を自前でやるにせよ、外注するにせよ、データ収集、データ分析の方法論についての知識が必要不可欠であることを示しています。これらの知識を教育・指導することも業務のうちに含まれています。

当然のことながら、組織的業務を営むためのさまざまなコミュニケーション・スキルも重視されます。したがって、ある程度の経験を積んだ人材でなければ、プログラム・マネージャーは務まりません。必須条件にプログラム評価に3年以上従事した経験が盛り込まれているのは、最低でもそのくらいの習熟期間を必要としている業務内容であるためです。

## 2) プラクティショナーのキャリアパス

科学技術政策の先進諸国において、一般的にプラクティショナーと呼ばれる評価専門人材はさまざまなバックグラウンドを持ちつつも、評価論に関する学習の機会を獲得し、評価マネジメントに携わっています。学習の機会がどのようなものであるかについては、研究開発マネジメントの発達に併せて変遷してきました。例えば、スウェーデンの研究開発資金配分機関である VINNOVA (スウェーデン・イノベーションシステム庁) の評価部門 (実質的な戦略形成部門) は分析能力を備えた高度人材が集積していますが、彼らのキャリアパスは世代によって異なっています (図 3-29 参照)。

VINNOVA のシニアクラスのマネージャー (もしくはディレクタークラス) は研究開発評価の創世記から歩んできた世代と言えます。技術者・研究者としてのキャリアから、オンザジョブトレーニング (OJT) 等によって、試行錯誤を繰り返しながら評価専門人材としてのキャリアへ転換を図りました。この世代の蓄積が、後に高等教育機関の専門コースの土台となっています。

ミドルクラスのマネージャーは、ある時期に職場を離れ、高等教育機関の専門コースで再学習し、評価専門人材としてのキャリア転換を図っています。専門コースでは、研究開発マ

マネジメント、経営学、経済学などの主に社会科学関連の修士以上の知識を習得しています。

もっとも若手のマネージャーもしくはスタッフは、最初から社会科学関連の大学院を修了し、専門職として採用されています。彼らの課題は、科学技術に関するバックグラウンドがないために関連する分野に関する知識を OJT で学ばなければならないことですが、研究開発評価やマネジメントについての専門性は最初から身に付けています。

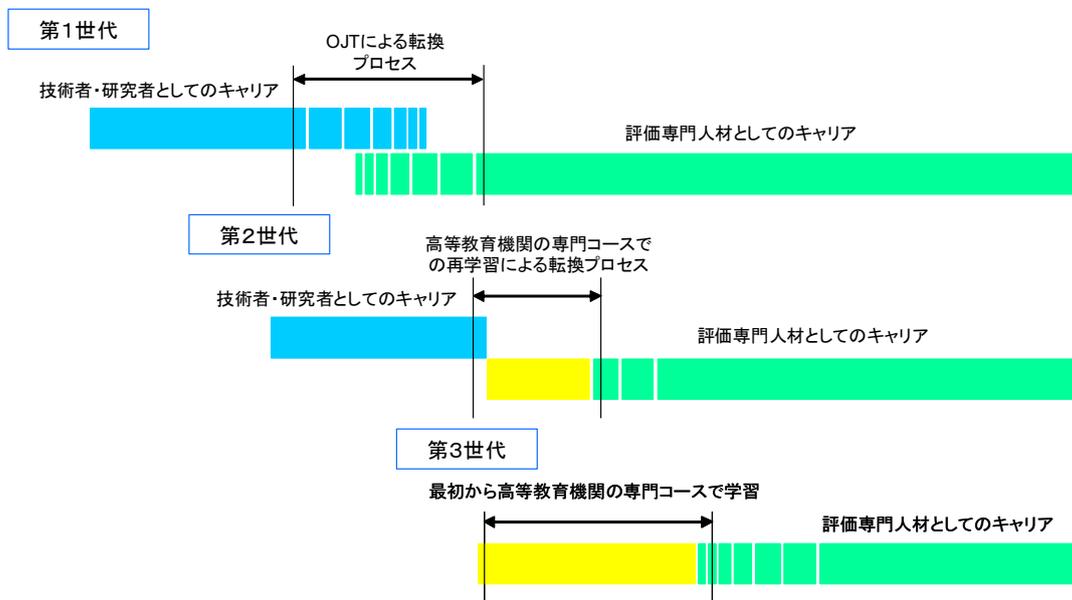


図 3-29 VINNOVA におけるプラクティショナーのキャリア形成パターン

出典: 財団法人政策科学研究所作成資料

また、学習の機会として現在でも非常に有益とされているのは、各国のプラクティショナーが集う研究開発評価や科学技術イノベーション政策に関する国際会議への出席と、そうした場における情報交換です。欧州やアメリカを中心に、このような国際会議が年複数回開催されているため、人々は積極的に参加し、研究開発評価のネットワーク作りを行っています。

我が国においても、平成 15 年には総合科学技術会議（現・総合科学技術・イノベーション会議）で評価専門人材の育成と蓄積が急務であることが議論されました（「平成 16 年度の科学技術に関する予算、人材等の資源配分の基本的方針」（平成 15 年 6 月 19 日））。研究開発評価を効果的に推進するためには評価を含めた研究開発マネジメントを担当する専門人材の育成こそが重視されなければなりません。マネジメントに関する知識や経験は外部から人を調達すればよいというものではなく、組織内部で自前で蓄積しなければ定着しません。その意味で、本テキストはプラクティショナーの入り口を示すための手引きである、ともいえます。

### (3) アナリスト

アナリストとは、研究開発評価の対象となるさまざまな事象を、専門的なツールを用いて分析することができる人材を言います。ある手法や技能についての専門性を有するだけでしたら、大学や研究機関の研究者もアナリストと呼べるかもしれませんが、研究開発評価の戦略、目的、枠組みの中で評価手法を位置づけ、全体を俯瞰しつつ分析もできるという点で、

より職業的能力を発揮する人材がイメージとなります。具体的には、研究開発評価を専門とする高等教育機関の研究者や公的機関の戦略形成部門の専門家、シンクタンクの専門家などが相当します。

アナリストはプログラム評価など複雑な対象の評価に貢献します。研究開発のインパクトを分析の対象とすることが多いため、社会科学関連のバックグラウンド（経済学、経営学、社会学等の PhD）を持った人材が主流です。彼らの駆使する分析手法は高度にシステムティックであり、データ収集、データ分析、システム分析などに関する知識が要求されます。また、評価の方法論だけではなく、政策装置（policy instruments）や科学技術の知識についても造詣がなくては研究開発評価を分析面から担うことはできません。

アナリストは、一定の専門性を身に付けた上で研究開発評価の支援を行うため、その育成方法は OJT によるものが主流となっています。

以下では、プログラム評価、プロジェクト評価、研究者の業績の評価、機関評価といった評価対象別に、評価のポイントをまとめます。

### 3.5 プログラム評価

#### 3.5.1 プログラム評価とは何か

2.2.3 において、「研究開発プログラム」は目的-手段の論理的な体系群であり、「研究開発プログラム評価」は第一義的な責任主体である政策立案者や研究開発を推進する側の行動及びその結果について評価を行うものである、と述べました。また、その目的は、プログラムの実施の可否を判断するとともに、研究開発の質の向上や運営改善、計画の見直し等につなげることやプログラムの改善、次のプログラム立案のための示唆を得るために行うものである、としました。さらに、プログラムを推進する上では仮説としての「道筋」が重要であり、その妥当性を検証し、その実施の可否を含めて、プログラムをよりよい方向へと導くための手段が「プログラム評価」である、とも述べました。

こうしたプログラム評価について、その目的や意義を事前評価、中間評価、事後評価、追跡評価というフェーズごとにまとめると、次のように表現できます。

表 3-14 評価のフェーズごとにみた評価の目的

フェーズ	概要
事前評価	取り組むべき問題やそれに対する他者を含めた取り組み状況を明らかにするとともに、プログラムの目的・目標自体が妥当かどうか、目的を達成するための諸活動が妥当かどうかの価値づけを行い、必要に応じて修正行動を行うこと。
中間評価	取り組むべき問題状況等の変化を捉え、プログラムの目的自体がなお妥当かどうか、目的を達成するための諸活動が機能しているかどうかの価値づけを行い、必要に応じて修正行動を行うこと。
事後評価 (直後)	目標が達成されたか、目標を達成するための諸活動が機能していたかについて価値づけを行い、次の活動等のために教訓を残すこと。
追跡評価	プログラムが存在していたことで、外部環境に対してどのような影響をもたらしたのかについて価値づけを行い、次の活動やよりメタな目的を達成するための手段の改善(政策提言等)に活用すること。

出典: 未来工学研究所作成

以下では、プログラム評価の基盤となる道筋を描く方法であるロジックモデルについて、詳細に解説します。

### 3.5.2 評価のポイント

#### (1) ロジックモデルとは

事前評価の段階で、有効な仮説としての「道筋」を描くために用いられる手法に、「ロジックモデル」と呼ばれるものがあります。

ロジックモデルは、C. H. Weiss が 1997 年にその著書の中で提唱したものであり、事業（プログラム）の構成要素を特定し、それらの要素間の関係性を描くとともに、活動がどのように成果につながっていくのかを線形のモデルとして表現したものです。これまで多様な分野でプログラムの実践的なマネジメントツールとして活用されてきました。科学技術イノベーション政策の領域においても多くの適用例があり、日本でも近年いろいろな主体に用いられるようになってきてます。ロジックモデルを用いることにより、プログラム推進により実現を目指すビジョン、そのために解決すべき課題、プログラム推進の方針、プログラム実施の具体的な取組、プログラムに影響する外的要因などが関係者間で「見える化」でき、共有することが可能となります。つまり、プログラム設計者の持つ仮説をみえる化するツールとして、また、関係者間のコミュニケーションツールとしてとても強力なものであると言えます。

また、プログラム推進の促進要因や阻害要因などが浮き彫りになり、PDCA マネジメントサイクルに基づいた変更や改善を関係者で共有しやすくなります。加えて、進捗や成果などを「見える化」するため、プログラム推進者以外のステークホルダーへ取組を説明し、理解を促し、共感を引き起こすことにも役立ちます。

図 3-30 は、図 2-2 で示した一般的なロジックモデルに、プログラム評価を行う上で重要となる 3 つの観点（①必要性・位置づけ、②プロセス・セオリー、③インパクト・セオリー）を足したものです。

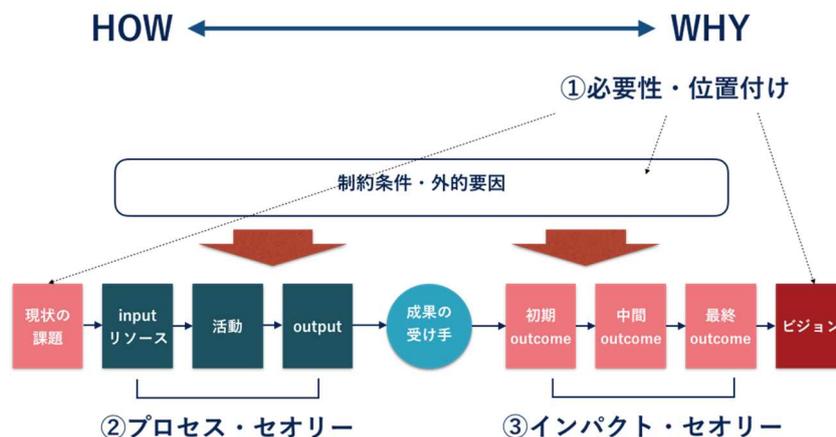


図 3-30 ロジックモデルと評価の観点

(出典) 林隆之氏資料をもとに作成

右端には、上位政策から与えられるプログラムが目指すビジョンがあり、左端には現状の課題をおきます。このギャップを埋めるものがプログラムです。こうしたプログラムを構想する上で重要となるのが、自然環境や世論・できごと、法制度や市場動向といった外的要因

であり、また、競合する他者やサービスの動きです。こうした外部環境の変化や制約条件を捉え、プログラム実施者の強みや弱みを考慮しながらプログラムを組み立てていく必要があります(①必要性・位置づけ)。

プログラムは、投入するリソースとそれを原資として行う活動(アクティビティ)群を通じて、何かしらのアウトプット(成果)を生み出していきます(②プロセス・セオリー。活動系とも言う)。一方、プログラムの活動や成果は、それ自体では価値を持ちません。ターゲットとなる受け手(カスタマー)に届いてはじめて意味を持ちます。プログラムの意図した結果がアウトカムであり、アウトプットはアウトカムの実現に向けた活動の水準を示すものと言えます。アウトカムは通常、いくつかの段階を経て、目指すビジョンに近づいていきます(③インパクト・セオリー。成果系とも言う)。いくつかの段階とは、たとえば、カスタマーがどのような意思決定や行為を行ったか(短期アウトカム)、その意思決定や行為を通じてビジョン実現に向けてどのような望ましい状況が生まれたか(中間アウトカム)、そして、意図した問題の解決が果たされたか(最終アウトカム)、といったものです。

このように、ロジックモデルは、ビジョンからの流れをさかのぼってみると(バックキャスト)、何が必要か、どのようにすべきかの説明になり(HOW)、現状の課題からみていくと(フロントキャスト)、なぜ事業が必要なのかの根拠を示すものとなっている(WHY)ことが分かります。

ロジックモデルは非常に強力なツールになりえますが、ロジックモデルのみでは、プログラムと価値実現との間の因果関係を十分に説明することはできないことに留意する必要があります。アウトカムが達成されたからといって、必ずしもプログラムがそれに貢献したとは言いきれないからです。ロジックモデルをプログラム設計や評価に用いる際には、アウトカムの達成に影響を与えるその他の外的要因や、アウトカムを達成するための補完的もしくは代替的な手段についてもあわせて考慮する必要があると言えます<sup>21</sup>。

---

<sup>21</sup> den Heyer (2002)は、ロジックモデルについて、1) 設計者や計画者が状況についての包括的な知識を持つ、2) 実施を阻害する外的要因はないという2つの非現実的な想定に基づくハードシステムとして批判し、経時的ロジックモデル(Temporal Logic Model: TLM)を提唱していますが、これは上級レベルの問題であり、ここでは詳細について触れません。

## 外部環境

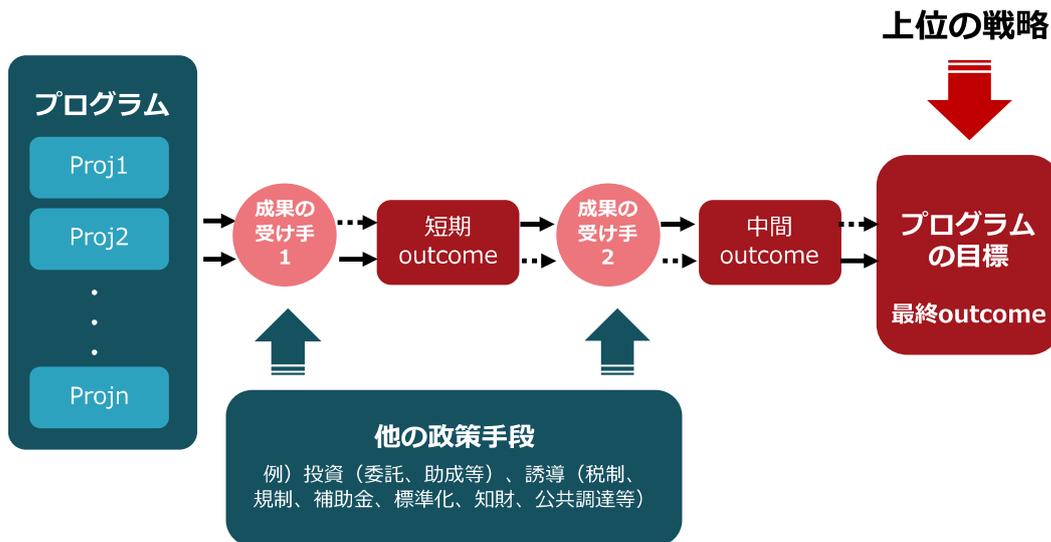


図 3-31 プログラム評価とロジックモデル

出典: 未来工学研究所作成

### (2) プログラムのデザインを確認する

プログラム評価においては、事前評価の段階において、ロジックモデルで表現されるような仮説をしっかりと設定、確認しておくことが重要性です。

一方、日本においては、政策分野や階層によらず、具体的な「アウトカム目標」やそこに至る「道筋」などの仮説（論理構造）が事前の段階で明確化されていることがこれまでほとんどありませんでした。つまり、そもそものプログラムにおいて、ロジックモデルのような事後的に検証可能な形で仮説が明確化されていません。したがって、次善のやり方として、調査分析や関係者との議論を通じて事後的に論理構造を明確化するなど、発見的方法によって仮説を再設定する必要があります。

ロジックモデルの作成は、本来プログラム推進者がステークホルダー等と一緒に進めるのが望ましいですが、中間・事後評価の段階ではすでに多様な諸活動が展開されており、それらを手掛かりに、まずはプログラム推進者が案を作成するのが効率的でしょう。ロジックモデルの具体的な作成方法については後述しますが、これはあくまで「次善」のやり方であり、本来はプログラム開始前のなるべく早い段階で、検証可能な形で用意しておくことが重要です。

### (3) 定期的なモニタリングシステムを構築する

前述のように、有効なプログラム評価を行うためには、中間、事後など評価の時期によらず、プロジェクトから情報を収集するための仕組みが必要です。プロジェクト関係者に過剰な負担をかけることなく、プロジェクトの改善を促すなどプロジェクト実施者に適度な緊張感とインセンティブを与えるものとなっていることが理想でしょう。

このための仕組みとして、2つの仕組みが考えられます。

まず 1 つは、プロジェクトの実績に関わるデータの収集システムです。実績には、論文や学会発表など研究の内容に関わるものと、シンポジウムやワークショップといった成果の社会実装に向けたイベントなどの諸活動に関わるものがあります。

中間評価の時期になってこうしたデータを集めはじめるとはプロジェクトの負担感も大きく、正確な情報が得られなくなる可能性があります。定期的に、もしくは逐次収集できる仕組みを適切に構築することで、成果創出に向けたプロジェクトの活動の活性化にもつながります。

もう 1 つはより本質的なものであり、プロジェクト実施者がみずからの活動や体制等を振り返り、必要に応じて改善につなげるための自己診断のメニューを提示することが考えられます。これは、各プロジェクトの実施者によるリフレクションを促進することを企図したものであり、マネジメント活動の一環としても位置付けられるものです。

### 3.5.3 ロジックモデルの作成方法

#### (1) 予備的作業

ここでは、プログラムの途上段階でロジックモデルを作成するような場面を想定し、解説します。これを行うためには、次の 3 つの観点からプログラムを捉えなおす予備的作業が必要です。

#### 1) 必要性の再設定

最初に行わなければならないことは、対象となるプログラムの必要性・位置づけ (Rationale) を定義することです。必要性の再設定を行うためには以下の作業が必要となります。

- 政策体系の整理
- プログラムの意義を裏付けるデータ (仮説の根拠)

政策体系の整理については、当該プログラムに関する上位施策からの位置づけと、類似プログラムとの差別化が必要になります。プログラムの意義を裏付けるデータについては、プログラムとして関心を持つ事象の特徴や推移を示すデータを用意し、プログラムが中長期的な目標としてどのような効果を実現するためのものかについての説明を行います。

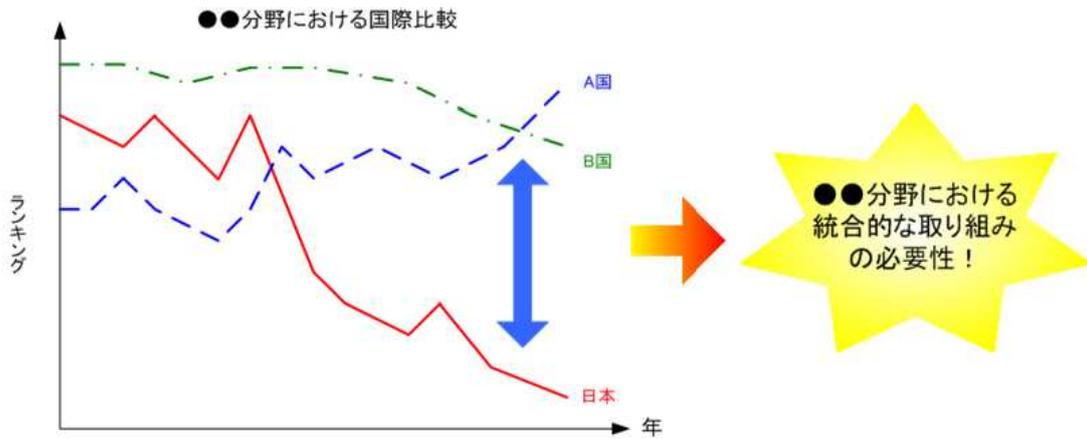


図 3-32 プログラムの意義の示し方

出典：財団法人政策科学研究所作成資料

## 2) 有効性の再設定

続いて、必要性のところでも述べた意義、つまり中長期的な目標とプログラム実施期間内に達成すべき短期的目標との因果関係や影響の大きさを考慮し、目的・目標を具体的に設定します。

ここで、陥りやすいのが次のようなロジック展開です。

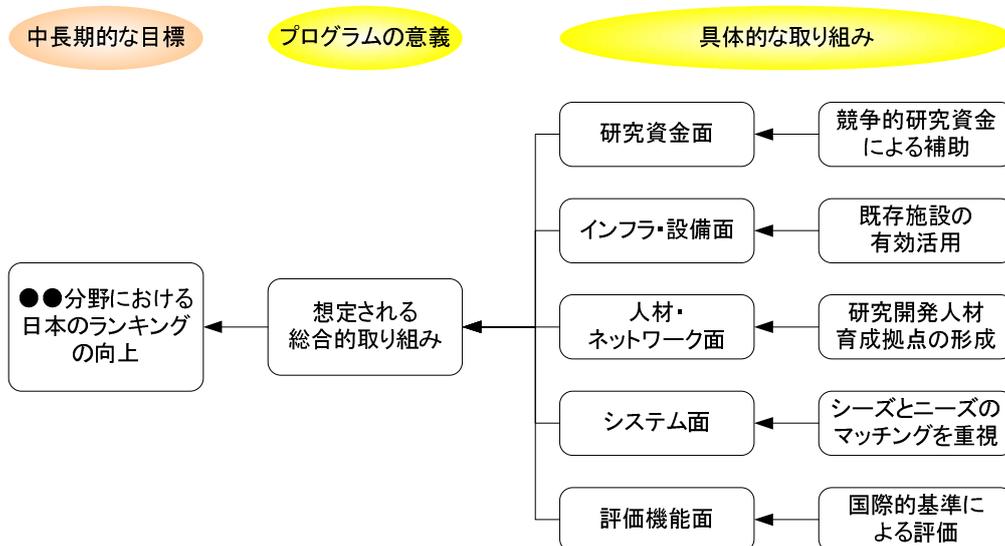


図 3-5 陥りやすいプログラムの設定の仕方

出典：財団法人政策科学研究所作成資料

中長期的な目標、つまり問題となっている状況の改善の方向を示したとして、プログラムによる総合的な取り組みの必要性を訴えるまでは十分な根拠と問題認識を示すことができます。しかしながら、なぜ個別の取り組みが必要かについての考察がプログラム起案者の経験的な知識から設定されているだけであって、具体的な根拠がありません。

プログラムの再設計のために本来なすべきことは、課題となっている状況の分析です。上

の事例では、なぜ諸外国と比較して日本のランクが下がったのかに対する分析が必要とされています。諸外国の科学技術政策の動向、予算の重点化のトレンド、人材開発のための仕組み・工夫、良質な研究開発課題の採択の仕方など、参考とすべき点を踏まえているかどうかはまず抑えられていなければなりません。また、国内固有の条件として良質な成果を生み出すためのリソースが充足しているかどうかの検討も必要です。関連研究領域の研究者の数、学会の動向、ネットワークの形成状況、インフラ・設備面の充足状況など、どこにボトルネックがあるかについての分析も必要です。つまり、政策研究が不可欠であるということです。

このような分析を踏まえないければ、具体的な取り組みとして掲げた項目の達成目標が議論できません。つまり、中長期的な目標に対するプログラムによって改善されるべき個別の施策目標が議論できないだけでなく、いつまでにどの程度の改善を図るべきかの目標水準の議論ができなくなってしまうのです。

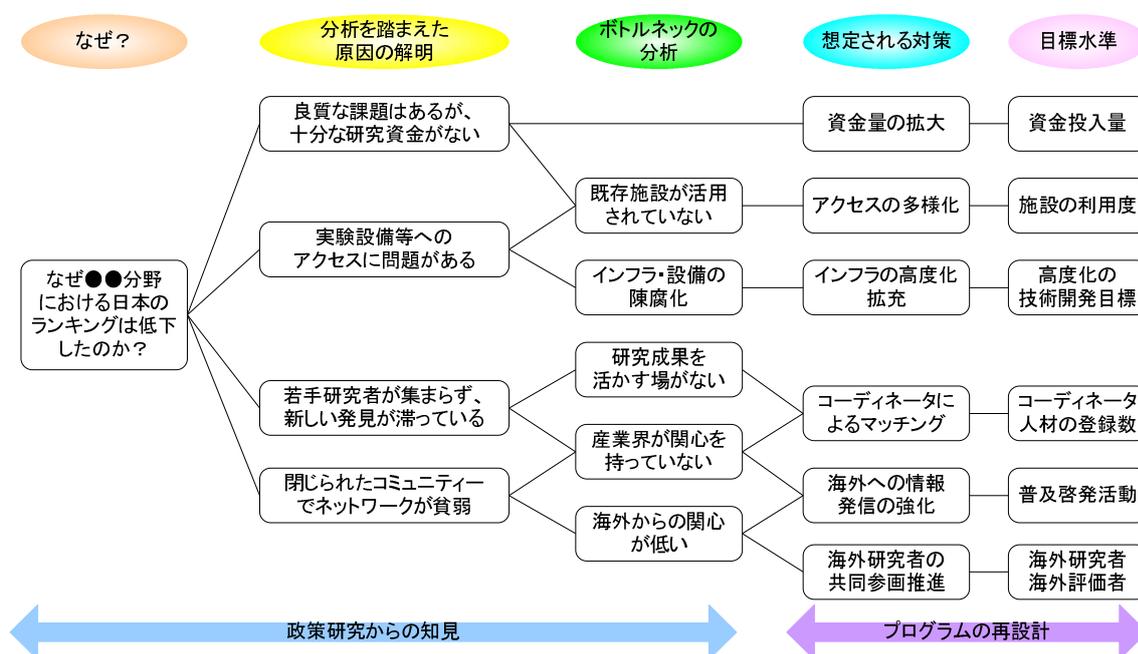


図 3-33 プログラムの有効性の再設定の仕方

出典：財団法人政策科学研究所作成資料

政策研究についてはプログラム実施機関が自ら調査・分析しても、適切な他の分析事例を参照しても構いません。審議会等の議論だけでなく、国や研究開発法人等が実施している数々の調査分析の報告書がこのようなロジック展開においてレファレンスとして活用されることで、政策がオープンに議論され、政策研究自体の質も向上すると考えられます。

さて、上のようなロジック展開を行うと何が見えてくるのでしょうか。中長期的な目標、つまりプログラムの目的は「●●分野における日本のランキングの向上」です。これは中長期的なアウトカムといってよいでしょう。次に、目標水準の項はプログラムの実施期間中であればアウトプットとして把握できます。原因やボトルネックの部分は、それぞれプログラムが達成すべきアウトカムといって良いでしょう。プログラムの所掌範囲はこのアウトプットとアウトカムであり、その結果を通じて果たして日本のランキングが上がるかどうかはプログラムのコントロールの範囲外ということになります。しかし、範囲内のアウトプット、ア

アウトカムについては確実に実現しなければなりません。

このようにプログラムの所掌範囲を明確にし、かつ実現すべきアウトカム（プログラムの目標）を論理的に定めることが、目的・目標を明らかにするという行為なのです。

こうしたロジック展開の他に、政策の有効性を示す方法としてはインパクトアセスメントと呼ばれる方法があります。インパクトアセスメントはある程度需要予測がつく政策領域に関して適用される方法論です。例えば、規制の影響、公共事業の経済性評価、教育プログラムや保健衛生プログラムなどの影響が相当します。インパクトアセスメントでは、プログラムの恩恵を受けるユーザーがどれだけいるか、ユーザーがどのような恩恵を受けるかを推計し、その影響（インパクト）の大きさをプログラムの有効性を明らかにします。

研究開発プログラムの場合でも、市場化や普及啓発までを視野に入れたプログラムにおいては、有効性のロジック展開をインパクトアセスメントの手法にしたがって行う必要があります。

### 3) 効率性の再設定

プログラムの意義、目的・目標が定まったとして、それをどのように展開するかは政策手段としてのプログラムの中身を明らかにしなければなりません。今度はプログラムの設計図を作る作業に移ります。これがなければ、個別課題の採択（Appraisal）や情報収集（Monitoring）、中間・事後評価（Evaluation）ができません。

プログラムの設計図に必要なのは次の要素です。

- 時間軸
- リソース
- アクター
- メカニズム

時間軸については説明の必要がないでしょう。研究開発プログラムの多くは予算執行上、事業の形態をとるために3年ないしは5年といった一定予算年度の枠が最初から決まっています。より詳細なタイムテーブルはプログラムの詳細が設計された後に実現可能なスケジュールを検討することで作成されます。

リソースについてはさまざまなものが考えられます。設計しようとしているプログラムの事業予算だけではなく、研究開発に携わろうとしている多様な参加主体のリソースが投入されてはじめてプログラムが実施可能となります。例えば地域における研究開発拠点の形成を目的としたプログラムでは、大学、企業、公設試験機関などがどのようなリソースを持ち寄り、持続可能な拠点形成メカニズムを提案するかが問われることとなります。ここでのリソースとはポテンシャル（潜在的能力）といっても良いでしょう。

アクターとはプログラムの枠の中で活動する組織や個人を指します。プログラムがどのような参加主体による活動を経て、最終的な裨益者（beneficiary）にどのような効果をもたらすのかを特定しなければなりません。

メカニズムとは、いかなるアクターのどのようなリソースを用いてどのような活動を経てアウトプットやアウトカムをもたらすかという、プログラムの仕組みのことです。

以上の要素を整理して、次のような大まかなプログラムのロジックモデルを考えてみましょう。

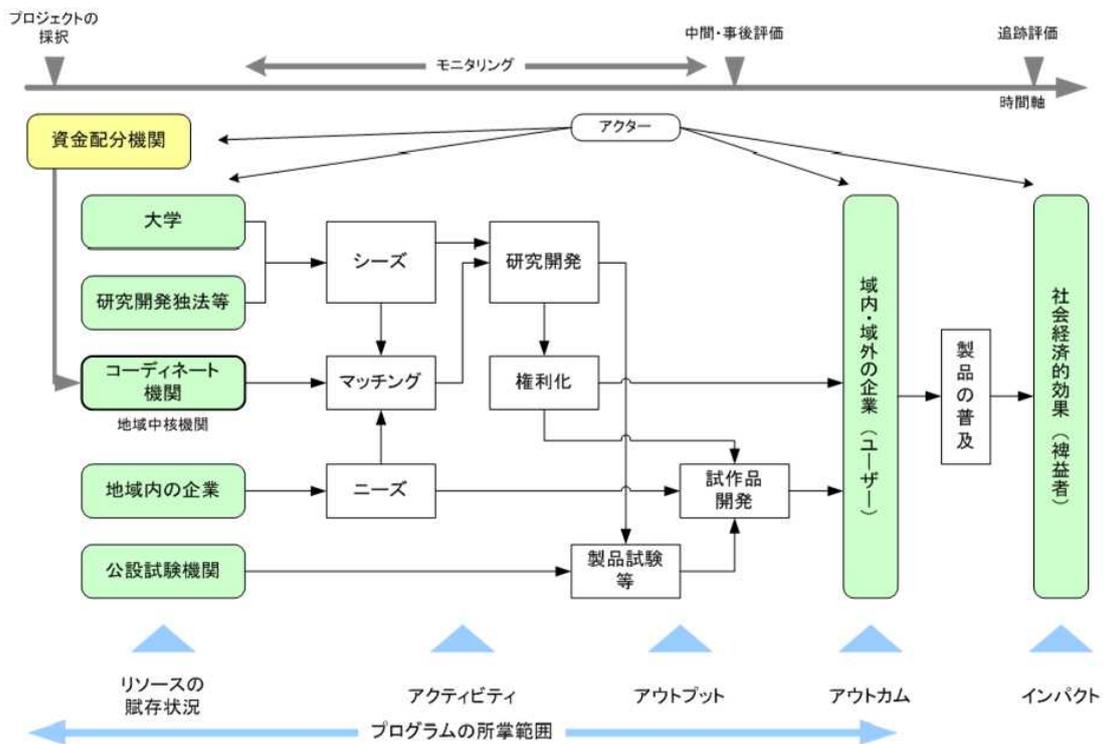


図 3-34 プログラムの時間軸・リソース・アクター・メカニズム

出典: 財団法人政策科学研究所作成資料

図 3-34 はある特定の技術開発分野に関する拠点形成プログラムをイメージしたものです。参加するアクターは大学、研究開発独法、コーディネート機関 (TLO 等)、地域内の企業、都道府県等の公設試験機関です。さらに研究開発の成果のユーザーとしては、域内・域外のプログラム非参加の企業、最終的な消費者が想定されます。プログラムではこのイメージに適合的な地域オリジナルの仕組みの提案を受け、採択した場合には中核機関に対してプロジェクト予算を配分するものとします。

メカニズムとしては、TLO 等のコーディネート機関が地域内企業のニーズと大学・研究開発機関等のシーズをマッチングし、課題を特定した後に研究開発を促進させ、得られた研究開発成果の権利化を行います。同時に、公設試験機関等を利用し、実用化研究を同時に進め、企業に試作品の開発までを請け負わせます。権利化・実用化された技術は、地域内外の企業からの引き合いを受ける段階で付加価値化され、市場での普及を通じて一定の社会的効果をもたらします。これらの効果も重要ですが、拠点形成により、こうした仕組みを持続的に発展させていくことがこのプログラムの主たる目的になっているものとします。

ここで、プログラムの所掌範囲は、研究開発拠点から生み出された技術等が権利化され、ロイヤリティを生むまで、あるいは試作品・実用化技術等が企業の引き合いを受けるまでとなっています。

以上のような状況を想定しつつ、効率性の再設定を行ってみましょう。

まず、プロジェクト採択 (Appraisal) についてです。プロジェクトの採択に関わる評価では、プログラムの目的・目標を達成するに十分なポテンシャルを有しているかどうかということが最大のポイントです。ここで、リソースの賦存状況に関して精査が行われなければなりません。図の例であれば、リソースの賦存状況として確認すべき点は次のようなポイント

トが挙げられるでしょう。

- 研究開発能力（実績等）
- 地域内企業の特徴、産業基盤等
- コーディネート人材の適切性
- コーディネート機関の業務執行能力
- 利用可能な地域内のインフラ・設備等
- 中核機関におけるプロジェクト管理執行能力

次にプロジェクトの進捗状況の管理（**Monitoring**）について考えてみましょう。採択評価の折に、中核機関におけるプロジェクトの管理執行能力を条件として挙げているのは、必要な時に必要な情報としてプロジェクトのデータが得られることを期待してのことです。プログラム評価はプロジェクト評価よりも広い観点から評価されます。例えば、競争的研究資金の枠組みであれば、ポストドクターの採用状況などがプログラムレベルで集計できなければなりません。中核機関の管理能力が低ければ、プロジェクト予算でどこの研究室がポストドクターを何人採用したかどうかすら把握していないことでしょう。大学教員等のエフォート管理も重要です。

管理能力がしっかりしているということは、必要なデータがすぐに提出できるということと同義です。データの蓄積方法については何かしらの指導が必要だとしても、プロジェクト業務の遂行中にデータを蓄積可能な人員・体制を確保していなければ、管理能力に問題があるということになります。

モニタリングとは、資金配分機関とプロジェクト実施機関との間にあらかじめ定められた項目についての情報供出に関わる契約があって、業務遂行中にデータを蓄積し、報告すれば良いだけの評価手続きです。モニタリングデータの推移に資金配分側として問題を見出すようであれば、プログラムのアドバイザー・ボード等の意見を聴取し、必要に応じてプロジェクトの責任者に対して指導を行うことができます。

中間・事後評価（**Evaluation**）の段階ではどのような再設定が必要でしょうか。アウトプットについてはモニタリングで情報を蓄積していれば、特に追加的にデータの提出を要請する必要はありません。上の事例であれば、アウトカムについてもデータ化することが可能でしょう（ライセンス件数、技術のオフナー数等）。プロジェクトから得られるこれらのデータを集計するために、プログラムレベルで集約可能な類型化が必要となります。

また、拠点形成プログラムのような取り組みでは、その仕組み（システム）が持続的かどうかという点が重視されなければなりません。拠点が形成されたことにより、マッチングファンドによる企業との共同開発が進んだ、優れた研究開発が他の競争的資金の発展課題として採択された、というような事例があれば、それらを類型化して他のプロジェクトと共通に実績を評価する必要があります。地域におけるネットワークの活性化を評価する場合でも、活性化を示す指標を他のプロジェクトと共通で指定する必要があります。

このように、プログラムの評価では個別プロジェクトから得られるデータをプログラムレベルで集計するために、またポートフォリオのようなかたちで分析するために、データの類型化（カテゴライズ）が必要になります。この作業は、本来、プロジェクトの成果報告書を作成する際に共通の仕様を課していれば済む問題ですが、プログラム評価よりも前にプロジェクト評価を先に実施している場合には、プロジェクトにより提供されているデータの量、範囲等が異なっている場合がほとんどです。こうした場合には、プロジェクト評価書のメタ

評価(共通の枠組みで比較すること)を前提に、データを収集しなおさなければなりません。

## (2) ロジックモデルの具体的作成手順

さて、準備が整ったところで、具体的にロジックモデルを作成する作業に入ります。ロジックモデルを作成するためには、次の構成要素を明らかにする必要があります。

- 活動 (activities) : そのプログラムの下で、関係するアクターが従事する主要な活動は何か。すなわち、アウトカムの達成に貢献する主要な活動は何か。
- アウトプット (outputs) : 主要な活動のアウトプットは何か。すなわち、活動が着手されたことを示すものは何か。アウトプットはアウトカム実現のための活動の水準であり、目に見えて (tangible)、数えられる (countable) ものとして把握される。
- 短期的アウトカム (short-term outcomes) : 活動やアウトプットから生じる短期的アウトカムである。ロジックモデルにおいてアウトカムは、概して「増大した」とか「改善した」というようなアクションワードを持っており、活動やアウトプットの結果を示している。プログラムの所掌範囲として最低限責任をもつべき結果として把握される。
- 中期的アウトカム (mid-term outcomes) : 短期的アウトカムの連鎖における次のつながりは何か。逆に、プログラムの意義でもある長期的なアウトカムの実現に必要な条件は何か。これらのアウトカムは、中期的アウトカムであると解釈される。
- 長期的アウトカム (long-term outcomes) : プログラム等の最終的なアウトカムは何か。どうしてこれらの活動が行われているか。これらは、一般に実感されるより長い時間がかかり、プログラム自体を超えた影響を仮定し、より戦略的なレベルにあることからである。

しかしながら、ロジックモデルにも記述しづらいプログラムの構成要素もあります。例えば、次のようなものです。

- どのようにプログラムが実施されるかについての、特定のマネジメント上の詳細事項
- 組織経営、設備運用、業務のプロセスに焦点をあてたもので、スタッフ雇用、装置の購入、ドキュメントのやり取りのような活動

これらは、重要なプログラムの活動ですが、通常、ロジックモデルに入れることはできません。

ロジックモデルを開発するためにはプログラムの主要な構成要素を特定する必要があります。一つの効果的なアプローチは、評価の専門家のサポートを受けながらグループディスカッションによってモデルを開発することです。プログラムに対してさまざまな視点を持つ個人は、ロジックモデルの開発に貢献できます。ロジックモデルに関する経験を持つ評価の専門家とプログラム・マネージャーの対象分野に関する専門知識を組み合わせることにより、ロジックモデルをより効果的に構築することができます。

さて、最初に議論すべきことは、下の表 3-15 のように、活動、アウトプット、アウトカム (短期的～長期的) に該当する要素をリストアップすることです。

表 3-15 ロジックモデル作成の第1段階

	活動 (activities)	アウトプット (outputs)	短期的アウトカム (short-term outcomes)	中期的アウトカム (mid-term outcomes)	長期的アウトカム (long-term outcomes)
構成要素	<ul style="list-style-type: none"> <li>既存施設の利用開放</li> <li>インフラ・設備の高度化のための技術開発</li> <li>科学技術コーディネータによるシーズとニーズのマッチング</li> <li>マッチング課題の発展的研究開発の実施</li> <li>各種普及啓発活動</li> <li>海外共同研究者の参画</li> <li>海外評価者によるパネル評価</li> </ul>	<ul style="list-style-type: none"> <li>施設の利用者増加、利用者属性の多様化</li> <li>インフラ・設備の高度化達成水準</li> <li>科学技術コーディネータによるマッチング件数</li> <li>マッチング課題の研究開発成果</li> <li>普及啓発の活動水準</li> <li>若手研究者の参加数</li> <li>海外共同研究者の参加数</li> <li>海外評価者の参加件数</li> </ul>	<ul style="list-style-type: none"> <li>既存施設のアクセスの多様化</li> <li>インフラ・設備の高度化・拡充</li> <li>産業界との共同研究の拡充</li> <li>海外からの成果引き合いの増加</li> </ul>	<ul style="list-style-type: none"> <li>良質な課題に対する十分な研究資金の助成</li> <li>優れた実験設備等へのアクセスの向上</li> <li>若手研究者の参入</li> <li>ナレッジの増加</li> <li>オープンなコミュニティの形成</li> </ul>	<ul style="list-style-type: none"> <li>●●●分野における日本のランキングの向上</li> </ul>

▲  
プログラムによる  
対策の活動内容

▲  
プログラムによる  
対策の活動水準

▲  
プログラムによる  
対策の結果

▲  
現在の状況の  
原因となっている  
個別要因の改善

▲  
改善したい状況  
(プログラムの意義)

出典：財団法人政策科学研究所作成資料

長期的アウトカムの欄については、プログラムの意義の部分で説明された、改善したい状況がそのまま入ります。中長期的アウトカムの欄については、関心となっている状況の分析から得られた、個別要因の改善目標が入ります。短期的アウトカムの欄については、それらにつながるプログラムが直接的に及ぼし得る対策の結果が入ります。ここまでの議論では、プログラムの有効性を明らかにした際の問題認識がそのまま適用できます。

プログラムのアウトプットについては、短期的アウトカムに寄与するであろう活動の水準が列挙されます。アウトプットは形式的には数値データとして把握可能なものとなります。活動の欄にはアウトプットを生み出す仕組みとして何を実施するのかの説明がリストアップされます。

次に、上の表 3-15 をフローチャートとして表現し、因果関係を考察していきます。

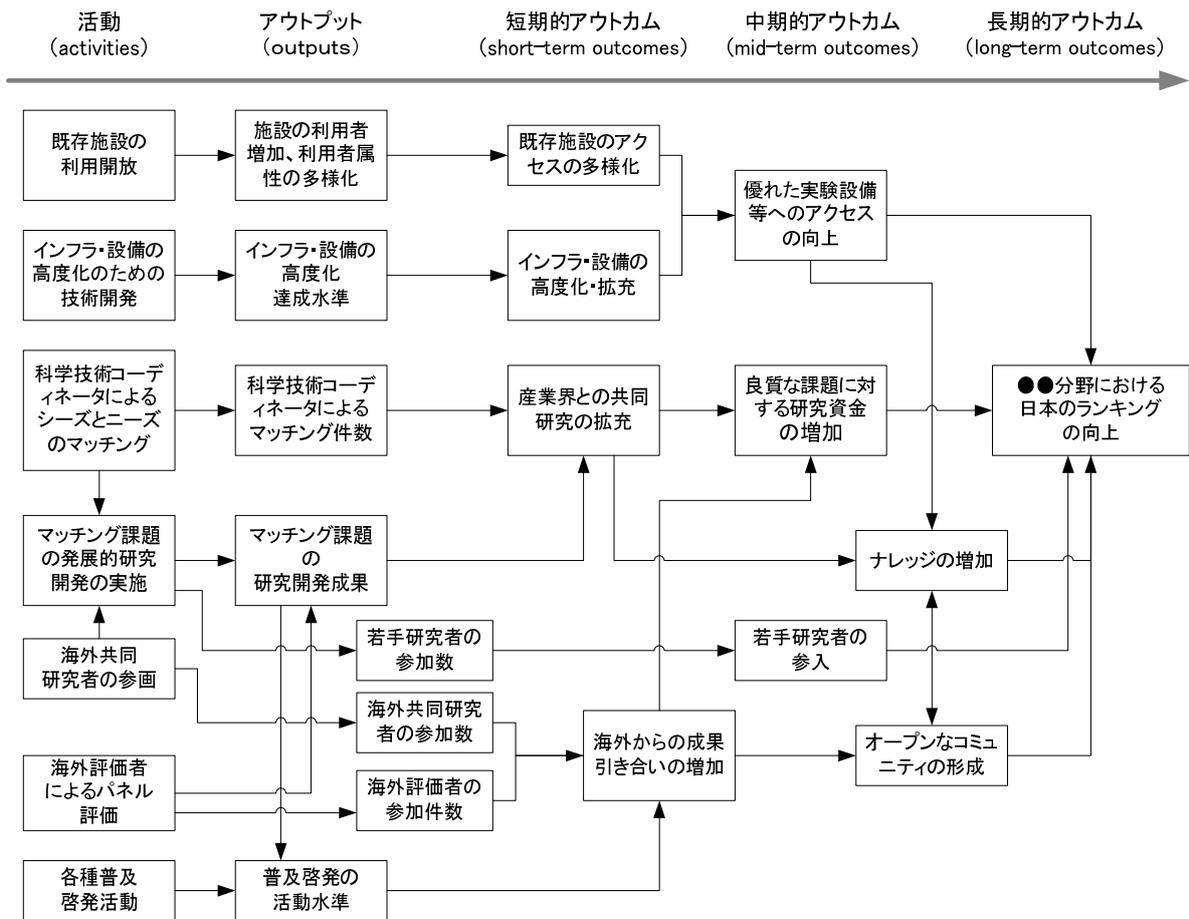


図 3-35 ロジックモデル作成の第2段階

出典：財団法人政策科学研究所作成資料

表 3-15 でリストアップした項目をボックスで配置し、それぞれの因果関係を矢印で結びます。原因となるボックスから結果となるボックスに対して→の矢印を引きます。上の図では、「ナレッジの増加」と「オープンなコミュニティの形成」が双方向の矢印で結ばれていますが、このような関係は共益関係です。お互いがお互いの実現に対して影響し合うような場合には双方向の矢印で結びます。

### (3) ロジックモデルのチェック

図 3-35 のような図がある程度完成した段階で、因果関係に誤りがないか、必要な要素が抜けていないかをチェックします。その際、可能であれば、評価の専門家に見てもらおうとよいでしょう。チェックは、特にアウトカムのロジック展開に関するフローに着目して行います。

アウトカムのロジック展開で重要な点は次の2点です。

- あるアウトカムは意図するアウトカムの直接的な因果関係を持つか？
- アウトカム同士の関係はいつでも成り立つか？

前者の質問に対する検討内容は、「アウトカムの実現に対して媒介的な機能（mediator）が抜け落ちていないか？」ということです。例えば上の図では、「産業界との共同研究の拡充」と「海外からの成果引き合いの増加」が「良質な課題に対する研究資金の増加」という中期的アウトカムをもたらしていますが、このような因果関係が直接的に成立するでしょうか。

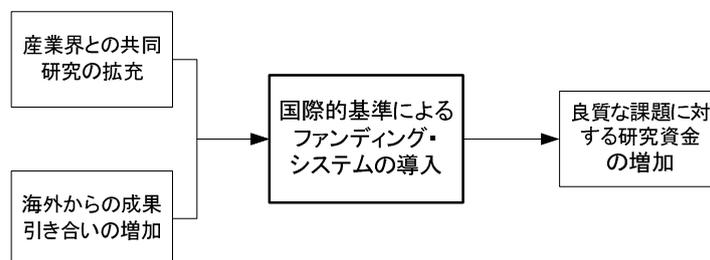


図 3-36 ロジックモデルにおける媒介的機能(mediator)の例

出典:財団法人政策科学研究所作成資料

直接的に、つまり放っておいても因果関係が成り立ちそうであれば、ロジック展開としては問題がありませんが、何かしらの媒介的機能がなければロジック展開に無理があると考えられる場合には、具体的な媒介的機能を検討することになります。上の例では、プログラムの次の措置として新たなファンディング・システムの導入の必要性を検討しています。

また、後者の質問に関する検討内容は、「アウトカムの実現に対して、社会経済的環境の変化が決定的に作用しないかどうか？」ということです。もし、アウトカムの実現に対して環境の変化が大ききようであれば、ロジックチャートにおいて、外的要因を調整的機能（moderator）として考慮する必要が出てきます。例えば作成したロジックモデルの図 3-11 で、「施設利用者の増加、利用者属性の多様化」と「既存施設のアクセスの多様化」の因果関係において、施設利用料金（コスト）が施設運営機関の事情で高騰した場合には、限られた研究開発実施者しか利用できないということになり、多様化が促進されるとは限りません。

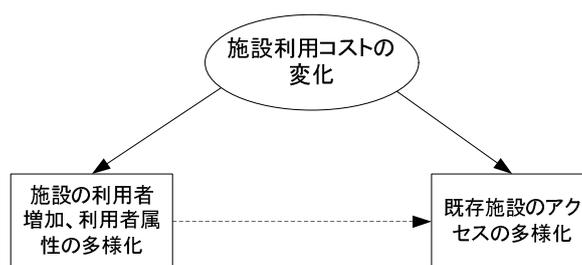


図 3-37 ロジックモデルにおける調整的機能(moderator)の例

出典:財団法人政策科学研究所作成資料

アウトカムが、環境要因が変化しないという前提で設定されている場合、政策担当者としてどこまでのリスク要因をプログラムの運営上考慮しているかという点がチェックの対象となります。

以上のような修正を経て、ロジックモデルは適宜見直される必要があります。

#### (4) パフォーマンス指標の設定

ロジックモデルでアウトカムを示した後に重要なことは、アウトカムが達成されているかどうかを判断するための評価指標を特定することです。指標の設定は非常に難しく感じられるかもしれませんが、上記までの段階で「言葉」としてアウトカムが明確に定義されているならば、評価として十分に機能するとも言えます。無理をして独自の指標を設定するのではなく、類似の評価事例を参考にしたり、評価の専門家のアドバイスをあおぎながら、より適切な評価指標を模索していけばよいでしょう。

ここでは、指標を設定する際の留意点について触れておきます。

まず、指標を設定する際に最初にぶつかるのが「何が定量的な評価になじみ、またはなじまないのか」といった疑問です。この線引きがどうあるべきかについては一概に言うことはできませんが、目標の達成度等を測る指標は、研究開発のリニアモデルやナショナル・イノベーション・システム<sup>22</sup>のような科学技術イノベーション研究に由来する概念的フレームワークに基づいていることを理解する必要があります。

また、どんなに優れたものであっても、指標は現実の完全かつ客観的な記述ではなく、現実を人工的に模した表現（**synthetic representation**）にすぎません。指標はそれにふりまわされるのではなく、意思決定支援のための道具として使用されるべきものであることを忘れてはいけません。

#### 3.5.4 評価の方法

以上みてきたように、ロジックモデルなどを用いて明確にした仮説をもとに、プログラムの目標達成を脅かしたり、促進する多様な要因をどのように情報として取り出すのか、プログラム目標の達成状況や達成に向けた活動の活性状況をどのように確認するのか、そして、それをどのように意思決定や資源配分の改善につなげていくのかを考えていくことが、プログラム評価をデザインするという事にほかなりません。

評価を行うにあたっては、様々な調査法、分析法、評価法を組み合わせる必要がありますが、プログラムの最低限の構成要素とはどのようなものでしょうか。評価のフェーズごとにどのような手法を用いるべきでしょうか。

---

<sup>22</sup> National Innovatoin System とは、個人、企業及び公的機関の間での技術と情報の流れが国レベルにおけるイノベーション・プロセスのカギであるとする考え方です。

表 3-16 は、オーストリアの研究・技術政策評価プラットフォームがまとめたプログラム評価の概要をもとに、評価のフェーズごとに利用可能な代表的な手法を足したものです。

表 3-16 プログラム評価の概要

	事前	中間／モニタリング	事後
質問	問題は計画されたプログラムを正当化するか？形成された目的は意図する方法で達成できるか？	どのようにプログラムを改善するか？プログラム・マネジメントを改善するものは何か？	プログラムは望んだ効果をもたらしたか？（継続されるべきか？もしそうなら、どのような修正がなされるべきか？）
評価の宛先	政策決定者	プログラム・マネジメント、（政策決定者）、プロジェクト	政策決定者、一般市民
特徴	概念的及びデザイン	形成的評価（学習への焦点）	総括的（フォローアッププログラムのための形成的評価）
チェックすべき点	関連性及びデザイン	プログラム実施の有効性及び効率性	有効性、効率性、達成された目的、有用性、プログラムの意図的及び非意図的効果
構成要素、方法	SWOT、戦略的概念のアセスメント、一貫性、ターゲット分析、コスト及び規模の見積もり、インパクト指標及び業績指標の一覧	組織、マネジメント及びプログラム実施のアセスメント、業績のチェック、継続に対する推薦	プログラムの全方位のアセスメント（内容及びプロセス）、質的、量的側面からのインパクト分析、他のプログラムとの比較（回顧的：経済的な長期のネットインパクト）
結果	プログラム・デザインの変更	組織及びプログラム・マネジメントの変更、焦点及びファイナンスの変更（適切な場合、プログラムの中止）	研究・技術プログラムの戦略的方向性（同一の政策分野における上位政策すべて）、フォローアッププログラムの構造及び目的の変更
最低限の構成要素	一貫性、戦略及びターゲット分析、インパクト及び業績指標	組織及びマネジメントアセスメント、プログラムの継続もしくは非継続の推薦	内容及びプロセスの全方位のアセスメント、インパクト分析
調査・分析・評価の方法	レビュー・パネル 事例分析（内外の類似プログラムとのベンチマーク等） ロジックモデルの設定 ステークホルダー分析 ステークホルダーへのヒアリングや協議 評価に必要なデータセットの準備を含めた評価計画の策定、等	レビュー・パネル ロジックモデルに基づく確認 終了したプロジェクトの追跡調査 プロジェクト関係者やプログラム・マネージャー等へのヒアリング、事例調査 ビブリオ分析、等	レビュー・パネル ロジックモデルに基づく仮説の検証 プロジェクトの追跡調査 プロジェクト関係者やプログラム・マネージャー等へのヒアリング、事例調査 幅広い利害関係者に対する調査 ビブリオ分析、等

出典：Platform Research and Technology Policy Evaluation (2012)をもとに未来工学研究所作成

これをながめると、いかにプログラムの事前評価が重要であるかが分かります。つまり、中間評価、事後評価は、事前評価の段階において設定した仮説（道筋）を検証し、必要に応じて修正を行ったり、教訓を残すために行われるものであるため、事前評価が十分に行われていないと場当たりな対応になりがちです。それにより、評価の信頼性を損なったり、評価に係るコストも非常に膨大なものになってしまいます。また、プログラムを構成する個々のプロジェクトも、どのような行動をとるのが目的適合的なのかが分からず、全体としての一貫性が失われてしまう要因にもなります。

なお、日本におけるプログラムには時限が設定されていることが多いですが、諸外国の多くではより長期もしくは時限の定めないプログラムがほとんどです。したがって、追跡評価は、プログラム単位というよりも大きな枠組み（テーマ横断的な評価や、機関全体の評価）の中で実施されています。また、継続しているプログラムにおいて、終了したプロジェクトの追跡調査（評価ではありません）は多くの機関でも行われており、プログラムの中間評価、

事後評価の中で活用されています。

### 3.6 プロジェクト評価

#### 3.6.1 プロジェクト評価とは何か

プロジェクトとは「ある一定の枠組み（＝プログラム）の下で展開される個別の研究開発課題」のことです。したがって、プロジェクトは、その（研究開発）目的、研究開発領域・分野等の基本的な方向性が上位の枠組みであるプログラムによって規定されており、プロジェクトの評価の観点（必要性、有効性、効率性）もプログラムの目的に合致するかどうかで判断されます。

また、個別プロジェクトの評価では、研究開発方法、研究開発成果の質などの内容的側面をいかに充実させるかが中心的課題になります。そのためには評価だけではなく、プロジェクトの採択から最終的な評価までを一連のマネジメントとして捉え、より良い成果が生み出されるように工夫をすることが重要になります。

図 3-38 はプロジェクトのライフサイクルを描いたものです。

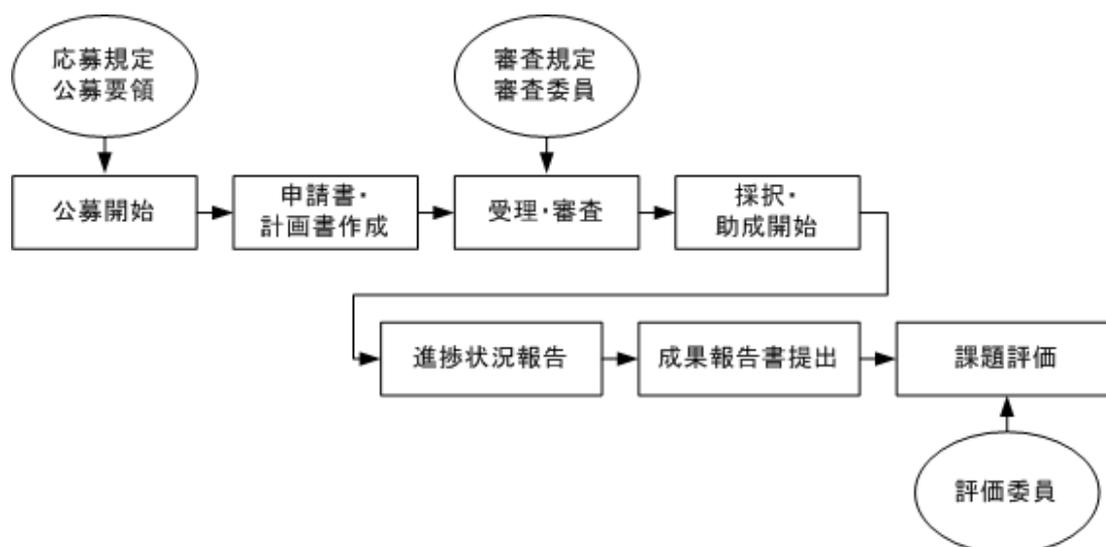


図 3-38 プロジェクトのライフサイクル

出典：財団法人政策科学研究所作成資料

通常、制度全体の応募規定や公募要領に従い、公募を開始します。研究開発実施者は申請書と研究計画書を作成し、公募期間内に提出します。プロジェクトの公募に際して、プログラム実施側が最低限行わなければならないことは、次のような点です。

- 公募要領の作成
- 採択評価の評価項目・評価基準の作成
- 公募方法・時期等の検討
- 採択評価の方法
- 採択結果の通知

このうち、公募要領の作成については、大筋のものはプログラム設計の段階で行われています。プログラムにおいて研究開発領域や課題設定が行われている場合には、それに対応した詳細な公募要領を作成する必要があるでしょう。採択評価の評価項目・評価基準は、研究開発領域毎、設定されているテーマ領域毎に作成される必要があります。これらの評価項目・評価基準については、公募要領等に記載して、申請者にも十分に通知しなければなりません。公募方法・時期等については、もっとも工夫の余地がある部分です。現状では、一部の資金配分機関において年複数回公募などの取り組みも行われていますが、さらなる工夫が考えられます。例えば、近年日本でもみられるようになりましたが、欧米の資金配分機関でよく行われているような2段階公募方式（A4で2枚程度のプレプロポーザルを提出・審査した後に、審査に通った申請者のみがフルプロポーザルを提出できる仕組み）や、通年で申請を受け付ける仕組み等があります。研究者の申請作業と制度実施側、ならびにレビューアの負荷を軽減しつつ、より良い研究課題を募集する仕組みについては、海外の資金配分機関に学ぶべき点が多いように思われます。

制度実施側は申請書・計画書を受理した後、審査に移ります。ここで、採択評価のためのマネジメントが発生します。レビュー方式、レビューアの選定が行われ、レビューが行われます。一方、制度実施側は資格審査や倫理規定に違反していないかどうかの判断をし、レビュー結果と付き合わせて最終的な採択の判断を行います。採択評価の方法はレビュー評価法が通常採られます。レビュー評価では適切な仕組みとマネジメントの構築が不可欠です。プロジェクトにおいてレビュー・システム、レビュー・マネジメントは中心的な要素になりますので、次項で詳細に見ていくことにします。採択結果の通知については、採択が決まった申請者に対してよりも、不採択となった申請者に対する配慮が必要です。我が国では、科学研究費助成事業の採択結果の通知システムが先進的な事例と言えます。

採択・助成が開始されたら、各種契約手続きが始まります。近年では、競争的研究資金の受給に伴い、研究者の所属機関に一定のオーバーヘッドを認めていますので、契約の対象は所属機関と結ぶことが一般的となってきました。不正使用問題などの教訓から、今後は所属機関側の管理体制がより一層、問われることになるでしょう。

研究開発がスタートしたら、モニタリングや途上評価を実施する必要があります。中間報告会などでアドバイザリー・ボードとの意見交換をするなどの機会を設けても良いでしょう。なお、近年では、テーマごとのマネジメントプロセスを複数のステージ（活動）に分割し、テーマとテーマの間に次ぎのステージに進めるかどうかを判断する関門（ゲート）を置く「ステージゲート方式」をプロジェクト評価の方法として採用するプログラムもみられるようになりました。これはプロジェクトの継続の可否を判断する査定的な評価であり、プログラム目的に合致する挑戦的な研究課題を複数走らせて様子を見る場合などに有効なやり方であると言えますが、たとえば科研費のような小型のプロジェクトについては、世界的にみてもこうした性格の評価を行うことは通常ありません。中間報告会などではむしろよりよい取り組みを促すための支援的な評価を行うことを目指すべきでしょう。また、プログラム評価の項目でも書きましたが、プログラムのマネジメント側がプロジェクト実施者側の自発的な改善を促す自己点検のためのフォームを提供するのも一案です。

研究開発が終了したら、研究開発実施者は成果報告書を提出し、これに基づいて課題の終了評価が行われます。採択評価と同様、レビューアを配置し、評価結果を導くためのマネジメントが必要となります。こうしたプロジェクト終了時の評価は、直接的にプロジェクトの改善に資するものではありません。その意味では説明責任を果たすための評価であると言え

ますが、誠実に研究開発を実施したにも関わらず芳しい成果がでなかった場合、その責任は誰に帰すべきでしょうか。プロジェクト実施者側に責任があるとされれば、萎縮して挑戦的な課題に取り組むことを躊躇してしまうようになるかもしれません。したがって、プロジェクトの終了時評価は、プロジェクト実施者側に次の取り組みの参考となる情報を提供するものであるべきであり、採択評価の仕組みやプロジェクト支援のあり方の改善など今後のプログラム・マネジメントへの教訓を導くために行うべきものと位置づけるのが妥当であると言えます。なお、小型・小額のプロジェクトについては、途上評価と同様、成果報告書などの提出を求めるにとどめ、終了時評価を行わないこともあります。

さらに、プロジェクト終了の数年後に、プロジェクトに対する追跡調査が行われることがあります。しかしながら、これはプログラム評価や機関評価などの一環として位置づけられるものであり、ここでは詳細について触れません。

このようなプロジェクトのライフサイクルを踏まえ、各段階でマネジメントを工夫したり、レビュー評価やモニタリングの仕組みを構築していく必要があります。

### 3.6.2 評価のポイント

ここでは、プロジェクト評価の中心となるレビュー評価の全体像を概観します。最適なレビュー評価というものではなく、実施するプログラムの目的・性格によって、さまざまな仕組みが考案されています。しかしながら、単に仕組みだけ作れば良いというものではなく、実務者のマネジメント努力が発揮されてはじめて有効なメカニズムとして機能します。

ここでは、レビューアをどのように選定するか、レビュー・システムをどのように構築するか、レビュー・マネジメントとはどのようなものかについて解説します。

#### (1) レビューア

前述のように、レビューアは専門家の立場からプロジェクトの内容について審査する評価人材です。ある特定のディシプリンの内部で評価する場合にはピアレビュー、複数の領域に跨って、あるいは新領域を評価する場合にはエキスパートレビュー方式が採用されます。

レビューアは組織や国から独立した存在であることが要請されます。つまり、個人の能力・資格によって評価に対して責任を全うするものであって、組織や国の代理人であってはなりません。非常にパブリックな存在であると言えます。

#### 1) レビューアの確保

レビューアは基本的にその国の研究者のストックに依存しますが、基礎研究や応用研究等の研究プログラムであれば、国外の研究者をレビューアとしてストックすることも可能です。例えば、EUが実施している研究開発プログラムでは、欧州域内の研究者をレビューアの対象としています。

レビューアの確保は、海外の資金配分機関でも悩みの種となっています。資金配分機関のプログラム・マネージャーは自分の担当する科学技術領域の学会誌に目を通し、国際学会や関連学会に頻繁に出席し、研究者の中からレビューアとして有望な人材をリストアップすることに余念がありません。例えば、全米科学財団（National Science Foundation : NSF）は毎年約 42,000 件の研究提案を受け付け、そのうち、審査によって高い評価を受けた約

11,000 件に対し研究資金を提供していますが、このような大量の申請を捌くためには常時 50,000 人の評価者を動員しています（1つの申請提案に対して、最低3人のレビューが審査に関わります）。このうち、毎年 10,000 人程度が申請審査を初めて経験するよう、リストの更新が継続的に図られています。

レビューをストックする機能を資金配分機関の外に配置している事例もあります。ドイツ学術振興会（DFG）では、ピアレビューの選定は DFG のプログラム・マネージャーが行いますが、ピアレビューのリスト化や能力保証についてはドイツ学術アカデミー連合などのサイエンスコミュニティが担っています。

研究者の立場からレビューを引き受けるとはどういうことでしょうか。先ほどの NSF の例ですと、多い人で年間 30 本程度のプレプロポーザルに目を通し、20 本近いフルプロポーザルを読み、かつパネル評価にも参加します。フルプロポーザルのレベルになると、学術論文を読むのと変わらないほどの労力を要します。

しかしながら、多くの研究者は NSF のレビュー依頼に対して、よほどの場合がない限り NO とは言いません。科学技術の振興、特に若手研究者育成のためという社会貢献の一環として評価への参画が社会的に認知されているからです。また、レビューができるということは研究者として一流であることの証でもあります。したがって必要な経費と若干の謝礼が支払われるとしても、多くの研究者は基本的にはボランティアでレビューを引き受けています。

我が国の場合をみると、政府全体の競争的資金の 5 割以上を占める最大規模の競争的資金プログラムである科学研究費助成事業では、第 1 段審査委員は 5,269 人です（平成 29 年度新規採択分）。審査員数も平成 11 年度と比べると約 5 倍に増え、一人あたりの審査件数も平成 11 年度に平均 183 件（最高 509 件）だったものが、平成 29 年度には平均 63 件（最高 150 件）と大幅に減ってきてはいますが、一方、レビューの質の確保という点では今後も試行錯誤を続けていくことになるでしょう。

## 2) レビューアの選定

日本の科研費や NSF の研究・教育プログラムでは審査委員が誰であるか分からないよう、匿名性を保持したブラインドレビューと呼ばれる審査方式が採られます。審査委員は研究領域毎にリスト化され、一定の申請数を割り当てられるように決めていきます。

一方で、申請者が自ら審査してもらいたい研究者を 1 名選択できるようなシステムもあります。UK の工学・自然科学研究会議（EPSRC）の産学共同研究プログラムでは、申請者が利害関係のない範囲で審査委員（レフリーと呼ばれます）を 1 名指名することが義務づけられています（その他に、EPSRC のプログラム・マネージャーが大学、研究機関等から 3 名のレフリーを選出します）。これは、先端的な研究領域という特徴もあり、申請者が自分の研究内容について、最先端の研究者に興味を持ってもらうための措置でもあります。

レビューは通常、1 課題につき 3～4 名程度が担当になります。論文の査読審査と同じ体制と考えてよいでしょう。各レビューアはプログラム実施側が用意する評価項目・評価基準に基づき、採択評価を行います。

### 3) レビュー・システム

レビュー・システムとは、プロポーザルに対するレビュー評価を採択審査の中でどのように位置づけ、結果を活用するかという仕組みのことです。

ここでは、欧州委員会 (EC) で実施されていたフレームワークプログラム (FP6) を例に、レビュー・システムを解説します。古い事例になりますが、基本形がよく示されています。

図 3-39 はプロジェクトの提案 (プロポーザル) の受理から採択 (融資の決定) に至るまでの手続をフローチャートにしたものです。

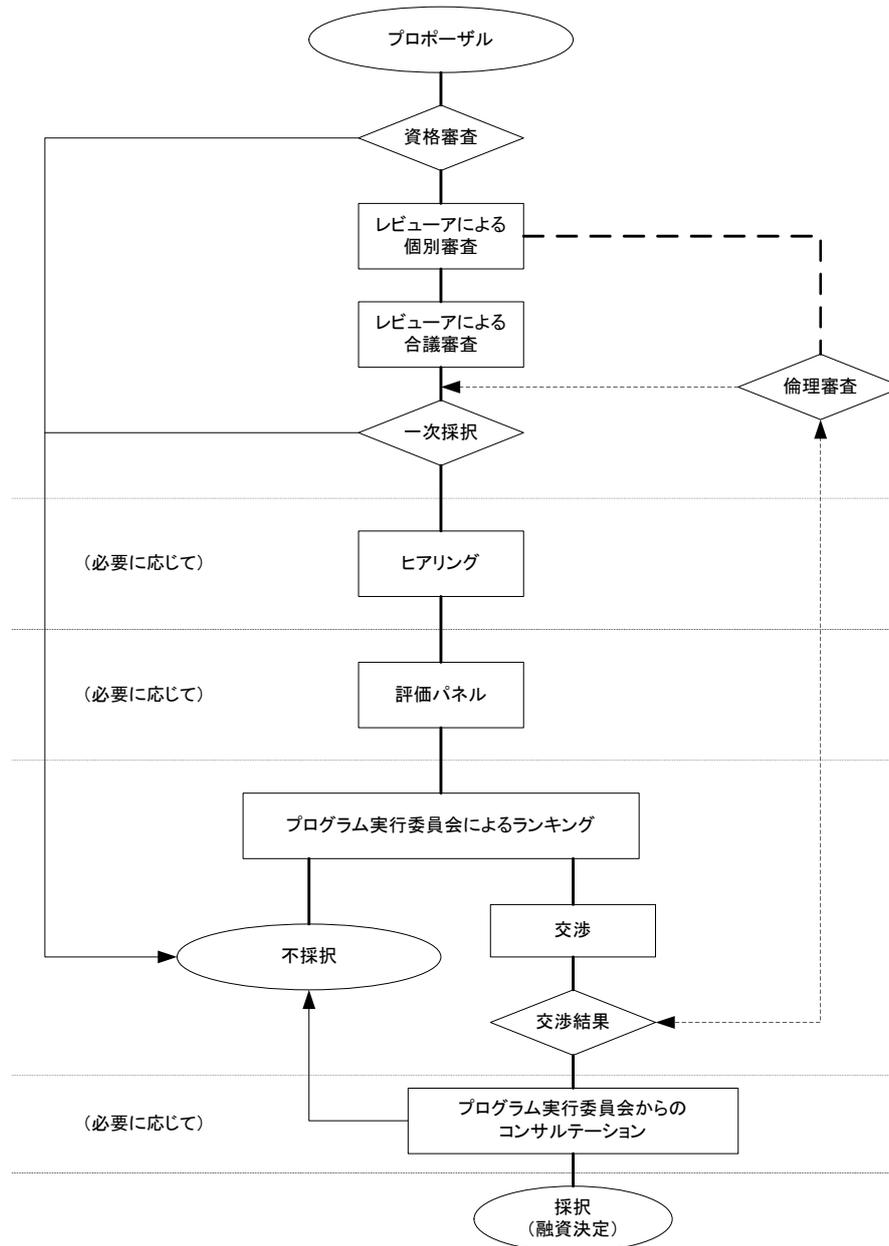


図 3-39 欧州フレームワークプログラム (FP6) のレビュー・システム

出典: <<http://www.cordis.lu/fp6/>>より財団法人政策科学研究所作成

プロポーザルが受理されると、まず資格審査に移ります。ここで FP6 の資格要件を満たし

ていないものは自動的に不採択になります。

資格審査を通りますと、サブプログラム（研究開発領域）毎のエキスパートが個人的に審査をします。審査はダブル・ブラインドレビュー方式（採択者・申請者双方が匿名）で行われます。FP6が実施されていた当時、エキスパートは欧州域内15ヶ国に点在していたため、メールレビューが基本でした。FP6では2006年の公募でおよそ3,400名のエキスパートが審査に関わっています。

表 3-17 FP6 のレビューアの数

プログラム分野	エキスパートの人数
1. Life sciences, genomics and biotechnology for health	603
2. Information society technologies	349
3. Nanotechnologies and nanosciences, knowledge-based multifunctional materials and new production processes and devices	31
4. Aeronautics and space	
5. Food quality and safety	191
6. Sustainable development, global change and ecosystems	400
7. Citizens and governance in a knowledge-based society	66
8. Policy support and anticipating scientific and technological needs	326
9. Horizontal research activities involving SMEs, research and innovation	※1
10. Specific measures in support of international cooperation	77
11. Support for the coordination of activities	※1
12. Support for the coherent development of policies	※2
13. Human resources and mobility, research infrastructures	1,149
14. Science and society	158
15. Euratom	33
<b>Total</b>	<b>3,383</b>

※1 2006年の応募実績なし

※2 “7. Citizens and governance in knowledge-based society”と共通

出典：<<http://cordis.europa.eu/fp6/experts2006.htm>>より財団法人政策科学研究所作成

個別審査が終了すると、同一課題に対して選出されているレビューア間でのパネル審査に移ります。ここで、一次採択としての評価結果をプログラムの実行委員会に提出します。

次に、オプション的な扱いになりますが、申請プロジェクトによってはヒアリング（オンサイトもしくはオフサイト）が実施され、一次審査のレビューアの他に数名のエキスパートを加えた評価パネルが構成されます。評価パネルはヒアリングの結果も踏まえ、最終的な評価点をプログラム実行委員会に告知します。

プログラム実行委員会は評価パネルが導いた各プロジェクトの評点をランキングし、採択候補を絞り込みます。その後、契約手続に移る段階で、申請者との細かい交渉が始まります。

また、FP6ではプログラムにおいて各種倫理規定が設けられています。プロジェクトの内容が倫理規定に抵触しないかどうかは採択評価のさまざまな段階で確認されます。

交渉の結果、プログラム実行委員会は申請者に対して、研究開発体制やプロジェクト管理体制、予算、倫理規定や支払規定に関するコンサルテーションを実施し、合意が得られれば、

融資が決定されます。すなわち、申請は最終的に採択されることとなります。

このように、採択評価の段階で、幾重にもエキスパートがレビューを行い、場合によってはヒアリングや現地見学等を実施するなどの工夫を重ねて、最終的な評価が行われます。ここまでの労力を採択評価に費やしているからこそ、その後の評価のプロセスがより合理的に実施されることにつながっています。

例えば、FP6 ではプロジェクトに関して中間評価のような制度化された評価は存在しません。FP6 のプロジェクトは3～5年の研究開発期間が与えられますが、各プロジェクトの実施者は、オープンな年次報告会等を開催し、広く意見聴取を行うなど、独自の活動によって自らの研究開発の進捗状況や研究開発の方向性を確認する機会を積極的に作っています。最終年度にはファイナルレポートを提出しますが、直後の事後評価も制度化されておらず、本格的な評価はプログラムレベルでの事後評価（終了後2年以内に実施）か、追跡評価（終了後5年以上の期間を置く）に委ねられます（ただし、財務状況の確認等、定型書式による年次報告は義務づけられています）。

## (2) レビュー・マネジメント

レビュー・システムを円滑に機能させるためには、さまざまな運営上の工夫が必要です。この工夫をレビュー・マネジメントと呼ぶことにしましょう。

レビュー・マネジメントはレビュー・システムの各段階で必要となります。ここでは、Kostoff (1997)を参考に、ピアレビューの質的向上のための8項目とそれに関連した留意事項を紹介したいと思います。

### 1) レビューアの所属機関に対する配慮

レビューアは大学や研究開発機関に所属する研究者であることがほとんどです。レビュー作業は、研究者に対して一定の労力を強いるため、プログラム・マネージャーがレビューアの所属する機関の上司に対して明確なコミットメントをしなければなりません。

コミットメントの内容は、レビューアに評価を依頼するプロポーザルの数、評価内容、委嘱期間、パネルとしての参加、パネルの開催回数・時間、交通費等の支給の有無など、通常の委員委嘱と同様の手続が明記されるべきでしょう。

また、レビュー作業が公益に資するものとしての認識をもってもらい、所属組織における研究者の評価を高めるよう、配慮しなければなりません。大学や独立行政法人のように交付金で運営されている機関に対しては、所属の研究者が国の実施する研究開発においてレビュー機能を提供しているという事実は、機関評価の重要な側面として認知されるべき点でもあります。

Kostoffはレビュー・マネジメントのもっとも重要な手続として、この項目を一番に挙げています。

### 2) 評価パネルの運営

評価パネルを設定した時に、評価実施側はパネルを上手にファシリテーションする必要があります。ピアパネルのレベルでは、1課題につき3人程度ですので、パネルリーダーを選出し、運営を任せるよりも、プログラム・マネージャーが直接ファシリテーションした方が

効率的であるとも言えます。

これに先立ち、プログラム・マネージャーはレビューアを選出する段階で、パネル評価の結果が偏らないように恣意的でない方法でレビューアを選ばなくてはなりません。

また、事後評価における評価委員会のような場においても、委員の質疑応答を先導し、彼らのコメントを要約し、評価レポートを作成するのは、ひとえにプログラム・マネージャーの力量に掛かっていると言っても過言ではないでしょう。

レビュー・システムが上手く機能するかどうかは、裏方であるプログラム・マネージャーのモチベーションと能力に掛かっていると **Kostoff** は指摘しています。

### 3) レビューアの能力と客観性

第3に重要な要素として、**Kostoff** はレビューアの能力と客観性を挙げています。ピアレビューアといえども、担当分野の専門的能力だけではなく、研究対象の複合的な側面、例えば隣接研究分野、研究により潜在的に影響を受ける技術・システム、プログラムのミッション等に対する理解がなければ、評価の役に立つことは出来ません。

一方で、プログラム・マネージャーもレビューアに対して、評価内容を専門分野に限定するのではなく、プログラムの目的や評価対象となるプロジェクトの最終的な目的にまで拡張して評価してもらうように配慮しなければなりません。このことによって、レビューアは評価経験を積み、革新的な提案についても公正な評価を行えるだけの潜在的な能力を身に付けることができます。

### 4) パネル間、分野間の正規化

レビュー・システムの採択判断に近い段階では、複数のテーマやプロジェクトを比較検討し、ランキングによって採択評価を行うなどの措置が採られます。そのため、評価パネル間、もしくは分野間の正規化を図るようにレビューアの構成を設計する必要があります。

ある程度類似性を持つ分野については、幅広いバックグラウンドを持つレビューアを共通に配置することで評価の正規化が図れます。極めて異なる分野については、プログラム実施機関の当該分野に関する戦略的価値を反映し、採点の厳しさや偏向性の補正が必要になります。その場合においても、レビューアのうち1名は幅広いバックグラウンドを持つエキスパートを配置し、評価内容の正規化を図るように努めるべきと、**Kostoff** は指摘しています。

### 5) 評価項目としての総合評価

レビュー評価を行う場合、設定された各評価項目の他に総合評価の項目を設けているケースが多く見られます。なぜ各評価項目の評点の和や平均点を用いず、別途、総合評価の評点を付けなければならないのでしょうか。

第1に、各評価項目に対するウェイトがレビューアによって異なるという点が指摘できます。これについては、より適切な評定区分の表現を工夫することで、ある程度の問題の解決にはなります。

第2に、より重要なことですが、プロポーザルを評価するための評価項目以外に、レビューアが重要と考える要素が必ず存在するからです。例えば、当該プロポーザルが国の研究開発よりも産業内の研究開発課題として相応しいと評価者によって判断された場合には、技術

的な評価項目の評点が高くても、総合評価の評点が低くなる可能性があります。

このような可能性がある限り、個別の評価項目・評価基準以外に総合評価の欄を設けることは有用と考えられます。

## 6) 評価の匿名性

レビュー評価を行う際に重要な要素として、評価の匿名性が挙げられます。通常、レビューアについては採択審査の期間中は匿名性が保持されます。一方、申請者のプロポーザルを匿名とすべきかどうかについては、プログラムの特徴によって判断されるべきでしょう。

例えば、学術研究のために広く助成されるようなプログラムでは、ダブル・ブラインドレビュー（申請者も匿名）が推奨されます。性別、国籍、年齢、所属機関名などに対する偏見を排除し、純粋にプロポーザルの中身について審査する方が学術研究の機会を提供するという面では大きなメリットになります。

一方で、応用研究や実用化研究などのプログラムの場合には、レビューアが評価項目として重視するのは、研究開発の実施体制や実績です。このようなプログラムでは、申請者を匿名とするメリットはあまりないでしょう。

## 7) ピアレビューのコスト

レビュー・システムは研究者がボランティアベースで参加する仕組みとはいえ、多大な時間コストの上に成り立っているシステムです。研究者がレビューに費やす時間は、1申請あたり1時間だとしても、割り当てられるプロポーザルが多ければ、とてつもない負荷になります。

また、申請する研究者もプロポーザルを作成するには、その何十倍もの時間を掛けていることでしょう。我が国の申請メカニズムのように、最初から各種申請書類と同時にフルプロポーザルを提出しなければならない場合には、研究者は1回の公募に全身全霊を掛けて望まなければなりません。たとえ他に申請できるプログラムの機会があったとしても、競争的研究資金制度においては重複申請が禁止されているために、審査結果が出るまでは申請することができません。完全なアイドル状態に陥るわけです。

そのため、公募型の研究開発プログラムが発達している諸外国では、2段階公募方式を採用している資金配分機関が少なくありません。1段階目では各種申請書類と簡単なプレプロポーザルを提出し、レビューアはプレプロポーザルの中から有望な研究開発課題をスクリーニングします。プレプロポーザルはせいぜい、A4で1～2枚程度に収まっていますので、割り当てられたプロポーザルの件数が多くても、時間コストを削減することができます。

一方、申請者の側も、プレプロポーザルの作成にはそれほど時間を掛けずともよく、さらに一次審査の結果が下りるまでは他のプログラムへの申請も同時並行的に可能です。フルプロポーザルを書くにあたっては、一次審査のレビューアからのコメントで、どこを詳しく書いて欲しいかの指示を受けていますので、より具体的なプロポーザルの作成が可能となっています。

こうしたレビュー・システムが成立するためには、十分な人員とマネジメントが不可欠といえます。

## 8) 評価の倫理基準

より良く設計されたレビュー・システム、配慮の行き届いたレビュー・マネジメントがあったとしても、最後に問題となるのはレビューアの倫理観です。

人が人を評価するということは、ある種の特権が与えられるということです。評価のプロセスにおいてはさまざまな利害の衝突が想定されます。またレビューアがその気になれば、秘密情報の流出、違法行為などの可能性がないとはいいきれません。

どのような文書にレビューアが署名したところで、最終的な拠り所になるのは、真の評価者たらんとする研究者の美意識・モラルです。それを支えているのは研究者コミュニティの自己規制しかありません。

レビューアが審査の段階ではなほだしい倫理規定の逸脱に及んだ場合、レビューアとしての資格取り消しや当該プログラムの申請停止処分といった対処だけではなく、学協会における会員資格の剥奪や職場での減棒等、厳しい対応が必要と考えられます。

また、それを行うに判断の拠り所となる倫理規定と罰則をあらかじめ作成しておくことが重要になります。

## 9) その他の留意事項

### ① パネリストの入れ替え

評価パネルの委員候補が固定化してしまうと、既得権益が生まれたり、評価組織が硬直化するなどの弊害が予想されます。一方で、すべての委員を毎回入れ替えてしまうと、評価の継続性がなくなり、レビューアの学習効果が発揮されなくなってしまいます。したがって、一般的にはパネリストの一部を毎年更新することが望ましいマネジメントになります。例えば、評価委員の任期が4年であれば、毎年4分の1を入れ替えるなどの工夫が考えられます。

### ② 利益相反（利害関係者の排除）

評価者がプロポーザルに対して利害関係を持つ場合、プロポーザルの審査から外れるよう配慮しなければなりません。しかしながら何ををもって利害関係と認めるべきでしょうか。

科学研究費助成事業では次のように利害関係を定義し、審査委員自らが該当すると判断した場合には当該研究課題の評価に加わらないこととしています。

表 3-18 科研費の利益相反の取扱

<p>(利害関係者の排除)</p> <p>第8条 評価に関する利害関係の排除の取扱いについては、次のとおりとする。</p> <p>一 科学研究費、特別研究員奨励費、国際共同研究加速基金(帰国発展研究)の場合</p> <p>(1) 評価者等自身が研究課題の研究代表者、研究分担者又は連携研究者である場合は、評価に加わらないこととする。</p> <p>(2) 評価者等が、研究課題の研究代表者、研究分担者又は連携研究者との関係において、次に掲げるものに該当すると自ら判断する場合は、評価に加わらないこととする。</p> <p>① 親族関係もしくはそれと同等の親密な個人的関係</p> <p>② 緊密な共同研究を行う関係 (例えば、共同プロジェクトの遂行、共著研究論文の執筆もしくは同一目的の研究會メンバーにおいて、緊密な関係にある者)</p> <p>③ 同一研究単位での所属関係(同一研究室の研究者等)</p> <p>④ 密接な師弟関係もしくは直接的な雇用関係</p> <p>⑤ 研究課題の採否又は評価が評価者等の直接的な利益につながると見なされるおそれのある対立的な関係もしくは競争関係</p>
--

出典：独立行政法人日本学術振興会「科学研究費助成事業における審査及び評価に関する規程」より抜粋

ここで重要なのは(2)の規定です。審査員がこの部分の規定に抵触するかもしれないということが直接的な問題なのではなく、審査員の倫理規範上、審査に関わるのが望ましくないと判断するかもしれないが問われています。例えば①について言えば、親族関係と同等の親密な個人関係とは、具体的にどのような関係なのかは人によって様々な解釈があり得るでしょう。しかしながら、親密な個人関係にあるから甘い評価をするということでないかぎり、本来は問題ないことです。

Kostoff が評価は倫理の問題であるといったのは、このように評価のメカニズムが個人の価値判断によって立つ部分が大きいからです。

### ③ 守秘義務

利益相反とならんで、評価者に課せられる倫理規定の一つが守秘義務です。採択評価の進捗状況や、成果物に関する情報等の漏洩は重大な守秘義務違反の例です。

例えば、アウトプットに関する守秘義務が守られないと、研究開発実施者側はその後評価に必要なデータをパネルに提出しなくなることが予想されます。企業の研究開発が絡む場合には、情報漏洩は重大な経済的損失を生む恐れがあるため、単にレビューアが罰則を適用されたというだけでは済まない問題となるでしょう。

形式的には守秘義務に関する誓約書(罰則等を含む)にサインをしてもらうという手続が必要ですが、重要なことは評価者の倫理規定として、科学技術コミュニティに広く認知してもらうということに尽きます。

### ④ オブザーバー

評価パネルにおいて、評価者、被評価者でもなく、事務局の側に立つ者でもない外部的な存在として機能する人材にオブザーバーがいます。オブザーバーの役割は、評価セッションの進行の仕方、公正性、公平性、手続を改善する方法や、評価項目・評価基準に対して独立した意見を述べることにあります。オブザーバーはプロポーザルや成果報告書に対する意見を述べるのではなく、評価という手続が公正かどうかを見る者です。

評価者も被評価者も、また評価事務局もオブザーバーの存在を意識することで、評価セッションを客観的に運営できることとなります。ある程度規模の大きい事業については、評価委員会を公開し、オブザーバーを配置するなどの工夫が必要となります。

### 3.6.3 レビュー評価の手続

#### (1) レビュー・システムの設計

最初に、プロジェクトのレビュー・システムを設計する必要があります。プロジェクトのプロポーザル提出から採択までのフローの中で、どのようにピアレビュー等を位置づけ、意思決定に役立てるかについて明らかにする必要があります。

ここでは、前節のFP6の採択評価のフロー図（図 3-39）を参考に、レビュー・システムのフロー図（図 3-40）を作成します。基本的な枠組みを先に作成し、オリジナルな試みについて加えていくことが頭の整理に役立ちます。

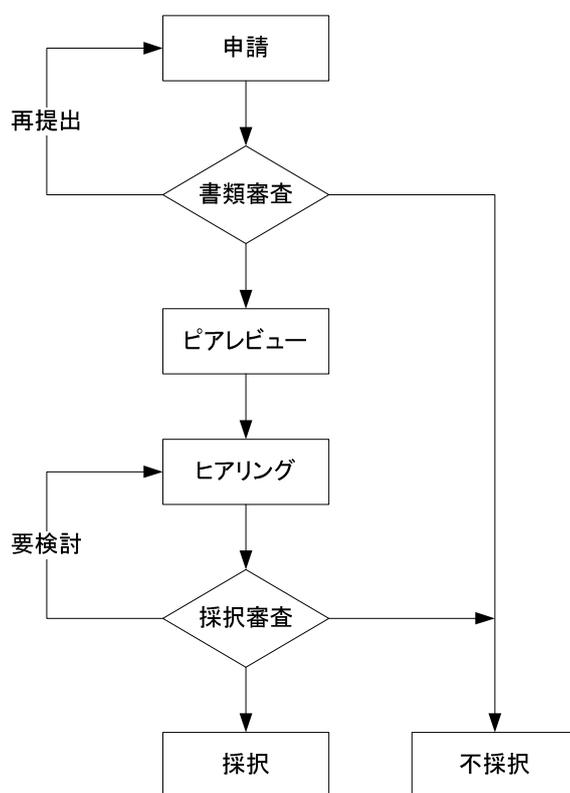


図 3-40 採択評価のフローの基本形

出典：財団法人政策科学研究所作成資料

図 3-40 にさらに必要な情報は、誰がどのような役割を果たすのかという責任分担に関する情報です。採択評価の責任分担を考えながら、実施体制を記載すると次のようなものになります。

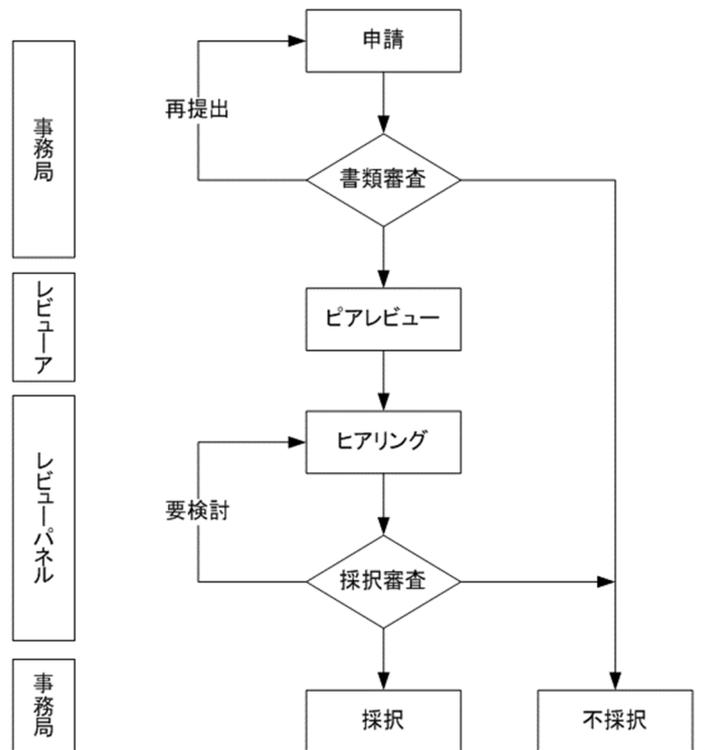


図 3-41 採択評価の手續と役割分担

出典:財団法人政策科学研究所作成資料

こうして作成したレビュー・システムを公募要領等に掲載し、申請者やレビューア等への採択審査のプロセスに対する理解を促す説明資料として役立てます。

## (2) レビューア・データベースの作成

評価に協力してくれるレビューアをどれだけストックしているかが、資金配分機関の資産といっても過言ではないでしょう。レビューアに関する情報はデータベースの形式で情報を蓄積し、資金配分機関内で共有することが一般的です。

データベースはフィールド（データ項目）の設計に情報収集の仕方、情報の活用の仕方がすべて反映されます。プログラム・マネージャーはレビューアに関する必要な情報がすぐに取り出せるよう、メンバー間で十分に協議した上でデータベースを設計しなければなりません。

表 3-19 では、専門分野等に関するデータとして、メニュー形式での入力を設定しています。資金配分機関がどのような専門分野のレビューアを欲しているかという観点から、活用のしやすい分野・領域の区分を設定し、レビューア候補者を分類していくことになります。

表 3-19 レビューア・データベースのフィールドの設計例

フィールド名	データ種別
氏名	テキスト
生年月日	日付データ
性別	2値データ
所属	テキスト
現住所	テキスト
Tel & Fax	リンク
e-mail	リンク
略歴	テキスト
主たる専門分野	メニュー形式
扱うことのできる隣接分野	メニュー形式
当機構における評価経験	テキスト
当機構のプログラム申請経験	テキスト
委嘱状等送付先住所・宛名等	テキスト
委員委嘱中	2値データ
備考欄	テキスト
謝金ランク	メニュー形式
謝金等先支払口座	テキスト
源泉徴収書等送付先住所	テキスト
支払履歴	日付データ
最終更新日	日付データ
最終更新者	テキスト

出典：財団法人政策科学研究所作成資料

作成したデータベースからピアレビューやレビュー・パネル等の審査員候補を抽出します。データベースはメンテナンスが重要であり、特に審査経験等のデータを反映することで、人材の拡充を資金配分機関が定型的な業務の一環として図っていくことが可能になります。例えば、毎年レビュー審査に関わる人材の1割を、レビュー経験のない新しい人材でまかなうという基本方針を固めることで、どの分野の人材がどれだけ必要となるかという具体的な情報収集戦略が立てられます。これにより、特定のレビューアに審査業務が集中しないよう、マネジメントができるようになるだけでなく、人材の層を厚くすることにも貢献します。

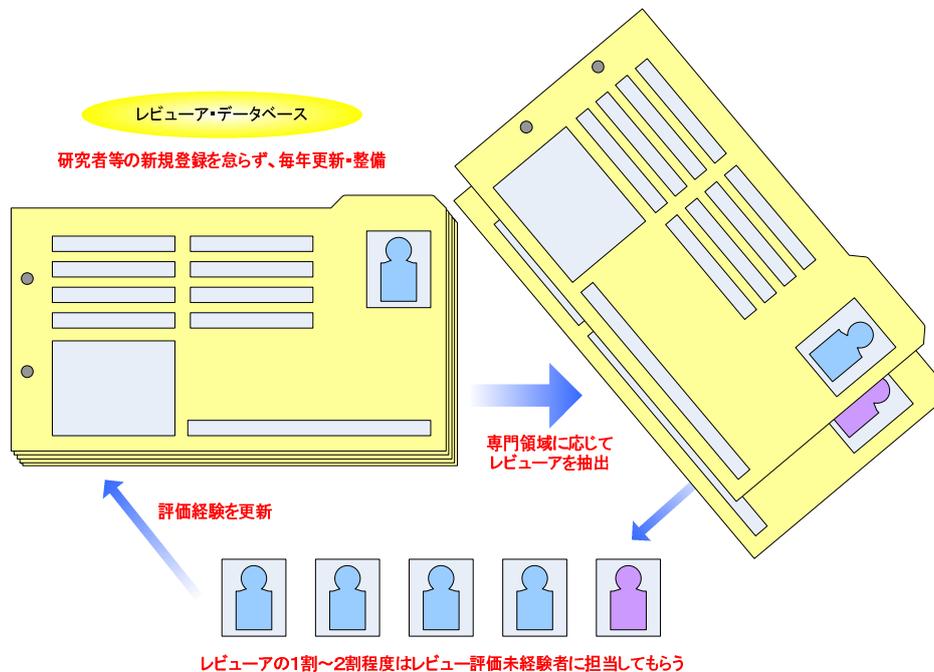


図 3-42 レビューア・データベースの活用と更新

出典：財団法人政策科学研究所作成資料

### (3) レビュー評価指針の作成

レビュー人材に審査を依頼する際に重要なことは、どのような審査をして欲しいかということをも具体的に示した文書を作成することです。

レビュー評価指針はおおよそ次のような構成を持ちます。

- 1) プロジェクトが属するプログラム、サブプログラムの目的  
プログラムの目的に則したプロジェクトの採択についての理解を促します。
- 2) 審査スケジュール  
レビュー評価の期日を定めます。
- 3) 利害関係者規定の説明と該当した場合の対処  
プロポーザルの申請者と利害関係にあると判断できた場合、審査を辞退してもらうこととなります。
- 4) レビュー審査の評価項目・評価基準・評定区分  
プロポーザルをどのような観点、基準から評価するべきかを明らかにします。
- 5) 総合評価点の評定区分  
総合的な評価点を付けてもらいます。また、その評価点の理由をコメントのかたちで記載してもらうことが必要です。
- 6) 倫理規定に関わる方針<sup>23</sup>  
プログラムに関係する科学倫理規定の適用を考慮して、倫理規定に抵触する恐れがあると判断される場合にはその旨を記載してもらいます(最終的な判断はプログラムの実行委員会等で決定されます。)

以上、6つの観点からみてきましたが、これらの中で我が国のレビュー評価システムの質の向上を図るためには、評定区分の定義をしっかりと示すことがもっとも重要です。

---

<sup>23</sup> 文部科学省であれば、科学技術・学術審議会生命倫理・安全部会の指針「ライフサイエンス分野における生命倫理に関する取組」〈[http://www.lifescience.mext.go.jp/bioethics/seimei\\_rinri.html](http://www.lifescience.mext.go.jp/bioethics/seimei_rinri.html)〉、厚生労働省であれば、「医学研究に関する指針」〈<http://www.mhlw.go.jp/general/seido/kousei/i-kenkyu/index.html>〉等が最新の倫理規定となる。

### 評定区分の悪い例

#### (1) 社会的・経営的ニーズ

- 5 : 基礎研究の技術は社会的・経営的ニーズとピッタリ合致している
- 4 : 基礎研究の技術は社会的・経営的ニーズと合致している
- 3 : 基礎研究の技術は社会的・経営的ニーズとある程度合致している
- 2 : 基礎研究の技術は社会的・経営的ニーズから若干外れている
- 1 : 基礎研究の技術は社会的・経営的ニーズから外れている

出典：長広仁蔵（1995）「評点法による研究開発の進め方と評価」日刊工業社。

上の評定区分の設定の仕方には2つの問題点があります。まず一つは、「社会的ニーズ」と「経営的ニーズ」という異なる評価の観点を同じ評価基準に同居させて評定区分を定めていることです。二つめには、「ピッタリ合致している」、「合致している」、「ある程度合致している」、「若干外れている」、「外れている」といった表現に対する判断基準が用意されていないことです。

このような評定区分の設定の仕方では、レビューアは「社会的ニーズはあるものの、経営的ニーズを満たしていない」場合や、合致している度合いについて何を基準にすれば良いかわからない場合には評点を付けることができません。

評定区分の設定には次の2つの原則を心掛けるようにして下さい。

- ・ 異なる評価基準を同じ文章に持ち込まない。
- ・ 優劣に対する判断基準を必ず文章中に記載する。

#### (4) レビューアへの委嘱手続

レビューア候補者は大学や公的研究機関、民間企業の研究所等、組織に所属している研究者がほとんどです。彼らに対して一定以上のエフォートを要求する以上、所属機関に対して正式な委嘱の手続を踏まえる必要があります。

委嘱状は次のような内容を備えている必要があります。

##### 1) 当該プログラムの審査員として委嘱する研究者への評価

レビューアとして相応しい経験・見識を有した候補者である旨を所属機関の上長に伝えます。

##### 2) 当該プログラムの審査に関わることの公益性に関する説明

所属機関側が研究者をレビューアとして参加させることの社会的意義を説明します。

##### 3) レビューアの負荷に関する情報

レビューアが何本のプロポーザルの審査を、いつまでに実施しなければならないかについての情報をコミットします。

##### 4) その他、謝礼・経費支払い等に関する情報

審査員に対して支払われる謝礼ならびに、必要経費についての情報を明確にします。

レビュー・システムは科学技術政策を振興する上で必要不可欠な仕組みです。審査員としてさまざまな組織の研究者が参加しやすいよう、研究者の所属機関に対して最大限の心配りをする必要があります。

## (5) 評価パネルの設計

メールレビューなどの評価方法だけでなく、パネル評価を実施する場合には、評価パネルの人数、構成、権限等に関する設計を行う必要があります。

プロジェクトの評価パネルの平均的な人数は5～7人程度です。評価パネルの人数は奇数で構成する方が、パネルリーダーの判断の責任を明確にすることにつながります。例えば、採択評価においてメンバー間で票が割れた場合、通常はパネル間での意見調整を行うことで票を集約させられることが考えられますが、結論が出ずにパネルリーダーに一任というかたちで最終的な判断を出す場合も想定できます。奇数で評価パネルを構成しておけば、最後の1票がパネルリーダーの責任として明確になります。

評価パネルの構成はプログラムの目的にも依存します。社会経済的なインパクトを目的に含むプログラムであれば、プロジェクトを評価するレビューアにはエキスパートが含まれていなければなりません。少なくとも、パネルリーダーは視野の広い、複合領域について扱えるエキスパートである必要があります。

また、評価パネルにまったく評価経験のない人材を最低1名は毎回組み入れるようにするとよいでしょう。評価の経験を通じて評価人材の育成・蓄積にもつながります。

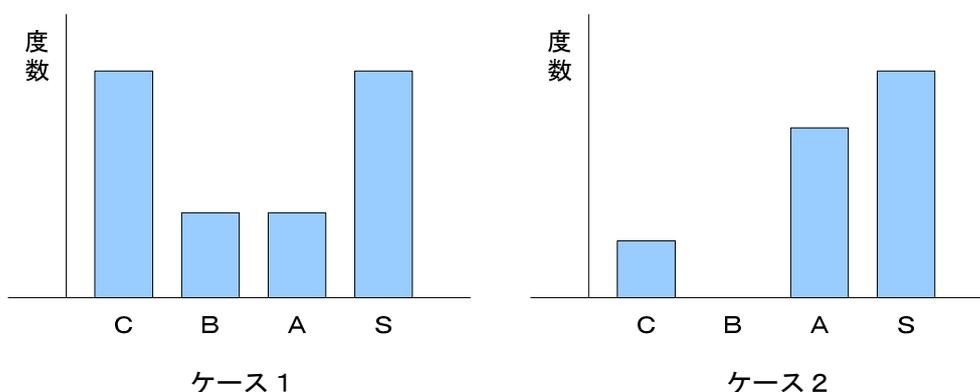
## (6) 採択の判定

プロジェクトの採択結果がまとまったら、プログラム運営主体は採択結果を速やかに申請者に知らせる必要があります。採択／不採択の結果は単純に総合評点（スコア）を知らせれば良いというわけではありません。なぜ、採択／不採択となったかの理由を文書にして明らかにするとともに、評価項目・評価基準毎の評点も明らかにすることが望ましい措置となります。

プロジェクトはプログラムが存続する限り、何度でも申請の機会が設けられています。また、一定数の申請者が存在しなければ、本来、競争的資金制度のような研究開発施策は成り立たない制度であると言えます。申請者にプログラムのリピーターとなってもらうためにも、とりわけ不採択の理由・評価結果の詳細情報を明らかにし、より良いプロポーザルを次の機会に申請してもらうことが、結果的にはプログラムの質を高めることにつながります。

## コラム 6 : 評価者（パネリスト）の間で評点が割れた時

ある研究開発事業の評価において評価パネルを活用して評点を付けたとします。この時、パネリストの付けた評点が次のような分布になったとします。



ケース 1 は評点が完全に割れた状態です。パネリストの間で当該事業に対する評価結果が正反対に出ているケースです。また、ケース 2 は少数のパネリストがまったく違う見解を示しています。このような場合、他のパネリストにはない独自の観点から評価意見を述べていることが想定できます。

いずれの場合においても、パネリストの間で意見調整を行う必要があるでしょう。新しい知見が得られる機会としてむしろ積極的に意見交換を図るべきです。なぜならば、一意の評価結果を決めることよりも、評価結果に至る理由を明らかにすることがより重要だからです。

時間的制約があるために、最終的にはパネル責任者（評価委員長）の権限で評価結果を一任するにしても、評点の集計には少なくとも 2 回のプロセスが設定されている必要があります。最終的にパネル間で合意に至らずとも、反対意見や少数意見はコメントのかたちで留意事項や課題として活かせることとなります。

## 3.7 研究者等の業績の評価

### 3.7.1 研究者等の業績の評価とは何か

研究者等の業績の評価は、2.2.5 で述べたように、当該研究者等を雇用する機関が評価実施主体となって行うものであり、「自己点検による意識改革、研究の質の向上、教育の質の向上、社会貢献の推進、組織運営の評価・改善のための資料収集、社会に対する説明責任等」を目的として実施されるものです。

現実としてこうした評価がどのような目的で行われているかについてみると、図 3-43 の通り、大学等については、国公私立によらず、「教員の自己点検による意識改革」、「研究の質の向上」、「教育の質の向上」といったメンタリング効果を狙ったものが多く、「教員の人員費削減」、「教員の業務負荷の平準化」、「教員の公平性や納得感の向上」といった人事考課に関わる目的で評価を行っているところは少ないように思われます。独立行政法人については、大学等と同様、「教員の自己点検による意識改革」、「研究の質の向上」が高い数値を示

しているほか、「教員個人に期待される職務内容の明確化」、「教員個人と組織の目標の連動」を目的とした評価も8割近い機関で実施されていることが分かります。

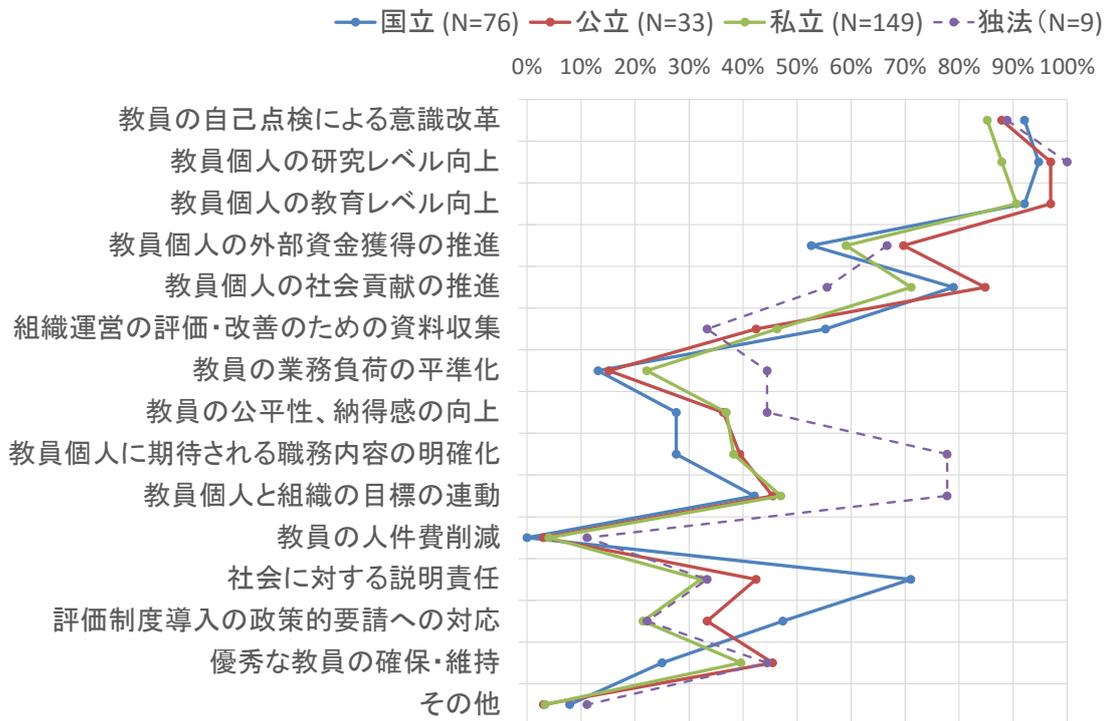


図 3-43 研究者（任期無し月給制）の個人業績評価制度を導入している目的

出典：林隆之，研究者等の業績評価，文部科学省研究開発評価人材育成研修(初級)，2017年11月30日。

### 3.7.2 評価のポイント

研究者個人の業績評価は、潜在的には人事考課とも密接に結びつきうるものであるため、文科省評価指針にも記されているように、「能力や業績の公正な評価の上、優れた努力に積極的に報いること等による公正で透明性の高い人事システムの徹底が重要」と言えるでしょう。

一方、研究者等の業績評価の結果が査定と直結すると、被評価者は「評価の論理」で行動するようになることが予想されます。研究者個人は評価結果が良くなるように、評価項目・評価基準に照らし合わせた活動を戦略的に取るようになるでしょう。この仕組みが上手く機能すれば良いのですが、研究開発の質の向上につながるような仕組みは評価項目・評価基準だけで最適化されることはあり得ず、日々のマネジメントの工夫によって達成されるものです。

したがって、多くの場合は研究開発評価の結果が個人の業績査定と直結すると、望ましくない影響をもたらします。組織構成員である研究者のさまざまな活動、チームとしての共同作業、ネットワーク形成機能、地域社会への貢献等を捨象し、ある特定の研究成果の面だけを査定の基準として取り上げると、評価を導入する前よりも組織文化の形成や研究者のモチベーションに悪影響を与えかねません。

研究者個人の評価については、研究開発評価とは切り離し、組織のミッションに従った公

正な人事評価を行うのが原則であると言えます。しかも、その考え方は過度に信賞必罰を反映した査定的なものではなく、真面目な研究者を奨励し、より良い成果を生み出すために支援的なものでなければなりません。

そのためには、組織のミッションに結びつく研究者個人の活動を列挙し、その枠組みの中での活動にどれだけ貢献(努力)したかを問うような多様な評価軸を構築するべきでしょう。何ら該当する活動がない、あるいは貢献(努力)の程度が低い人間がマイナス査定になるのであれば、正直者が報われることを担保することができます。

### 3.8 研究開発機関等の評価

#### 3.8.1 研究開発機関等の評価とは何か

2.2.6 において、大学全般については「学校教育法」に基づく自己点検・評価や認証評価が、国立大学法人及び大学共同利用機関法人については「国立大学法人法」に基づく法人評価(教育研究の状況についての評価を含む)が、研究開発法人等については「独立行政法人通則法」に基づく法人評価が義務づけられていることを述べました。

これらの評価の基本となるのは、機関自らが実施する評価であり、それぞれの機関の特性等に応じて、これを充実させていくことが期待されています。

また、機関評価は、前述のように、原理的には機関を1つのプログラムとして見立てた「プログラム評価」の一種という見方もできます。もちろん、機関評価は組織自体のあり方や戦略形成の問題とも関わってくるため、必ずしもプログラム評価と同じであるとは言えませんが、機関の行う諸活動が機関のミッションや中(長)期目標の達成に向けて適切にプログラム化されていないと、真に有効な機関評価は実施することができません。繰り返しになりますが、大学等にせよ国立研究開発法人等にせよ、機関評価において基本となるのはやはり機関の活動を構成する個々のプログラムを単位とした自己評価であり、プログラム評価を実質化することが組織経営上欠かせなくなってきました。

以下では、応用的な内容もたぶんに含まれますが、プログラム評価の枠組みを援用し、機関評価を行う際のポイントをまとめていきます。

#### 3.8.2 評価のポイント

ある組織を評価するといった場合、通常は定型的な枠組みに従ってある時点におけるデータを見ることになります。財務諸表が良い例でしょう。財務諸表は企業の一年間の経済活動に関する成績表です。企業は経理を中心とした経営内容を財務諸表のかたちで社会に公表し、利害関係者の監視を受けることで説明責任を果たすことができます。

財務諸表は企業の活動の結果ですから、活動に対する努力が実を結べば、自ら財務諸表上の経理指標は改善されることが期待されます。企業の利害関係者は結果としての数値を毎年見ること、企業の全体のパフォーマンスをチェックすることができます。

ところが、財務諸表のデータは企業の経営者にとって企業の健康状態を教えてくれても、どうすれば健康体になれるかについての情報を教えてくれません。例えば、ある業務を効率的に実施し、業績を伸ばすと同時にコストも下げるという目標に対して通常の企業会計データは役に立たないのです。つまり、ある業務を中心としたコストデータの管理ができなけれ

ば、業務の改善につながらないということです。

業務の改善のためには管理会計とよばれる会計手法を導入する必要があります。一般的な経理データから、特定の業務のために費やされた直接経費、間接経費を何からの基準を用いて配賦し、業務の原価を計算します。

管理会計手法を導入するためには、業務がどのような活動から構成されているかという業務分析、活動のために費やされるコストを計算するための方法論の開発、通常の間接データを必要に応じて読み換えることのできる経理システムの導入など、大変なマネジメント費用が発生します。しかし、それだけの投資を行うことで、ようやく業務の改善に対する評価が可能になります。そして、改善が結果として企業の財務諸表における経営指標の向上に結びつくのです。

翻って、国立研究開発法人等や大学等について考えると、前述のように様々な法体系のもとで機関評価が実施されています。たとえば、国立研究開発法人の場合、機関の設置目的（ミッション）から展開される具体的な活動領域と実施している業務の関係性を明らかにした上で、中長期計画期間の実施業務がどのようにミッションに貢献しているかを評価しなければなりません。

こうした機関評価で問題となっているのは、実施している研究開発事業等の評価と機関全体のパフォーマンスの評価を結びつける方法論がないままに評価の重複が起りかねない状況にあることです。具体的には次のような問題が挙げられるでしょう。

- 機関の設置目的（ミッション）と、それから展開される活動目的・目標に対して、実施している事業の目的・目標が直接的に結びついていない。（実施事業の評価結果を機関評価に利用できない）
- 実施事業ならびに機関全体のパフォーマンスを測定するための方法論が導入されていない。（業務分析が出来ていない）

こうした問題を考えるときに、先に紹介した管理会計的な考え方が大変役に立ちます。機関を評価するためには、その構成要素をどのように分類し、いかに再構築するかということが大変重要になります。

### 3.8.3 評価の内容

#### (1) 機関の何を評価しなければならないのか？

民間企業はあるビジネスモデルにしたがってキャッシュフローを確保することで経営が成り立っています。したがって民間企業を評価する場合には、財務的な指標を観ることで、実施している事業の効率性等を企業経営の総体として評価することができます。

ところが、大学等や国立研究開発法人等は多かれ少なかれ公的資金を運用して事業を展開しています。機関の依って立つ理由は、独立行政法人通則法や各独立行政法人の個別法、国立大学法人法等によって明らかにされているように、何らかの公共的なミッションを果たすことにあります。これらの機関についてはその財務状況をみて評価が可能なのではなく、機関のミッションに沿った個別具体的な事業展開について、その改善努力や事業の効果を国民の観点から評価しなければなりません。

つまり、公的機関の評価では機関の「状態 (state)」を診るだけでなく、機関の「活動 (activity)」についても診る必要があります、その「活動」は組織のミッションに結びついている必要があります。

## (2) ミッションから活動まで

国立研究開発法人の場合、法人の設置目的（ミッション）を果たすために、目標期間中に以下の点を明確に定める必要があります。

- 1) 中長期目標の期間（5年～7年）
- 2) 国民に対して提供するサービスその他の業務の質の向上に関する事項  
（国立研究開発法人の場合は研究開発の成果の最大化その他の業務の質の向上に関する事項）
- 3) 業務運営の効率化に関する事項
- 4) 財務内容の改善に関する事項
- 5) その他業務運営に関する重要事項

この中で、中長期目標を実現するための中長期計画を策定するに当たって記述に工夫が必要な事項は2)と3)です。なぜならば、当該法人のミッションと国民に提供するサービス（若しくは研究開発の成果）や業務を具体的に展開し、その上で個別業務の効率化について述べなければならないからです。さらに、これらの展開がきちんと実施されないと、毎年の年度計画が具体的に記述できないことになります。

法人のミッションから具体的な活動に至るまでの連鎖がしっかりと出来ていないと、中長期計画・年度計画は策定できません。そして、法人評価は中長期計画、年度計画の目標達成度を評価する仕組みになっているので、評価自体が成り立ちません。これは大学等の場合であっても同様です。

ミッションから活動に至るまでの連鎖を作り上げるということは組織の理想モデルを構築するという行為に相当します。

一方で、現状の法人の組織構造がどのようなシステムを形成しているか（組織の実態はどのようなになっているか？）を整理して、組織の実態モデルを用意する必要があります。組織の実態モデルは単に組織図や機構図を見れば良いというのではなく、組織内部の部署がどのような機能を（相互に）果たしているか、という観点から捉える必要があります。

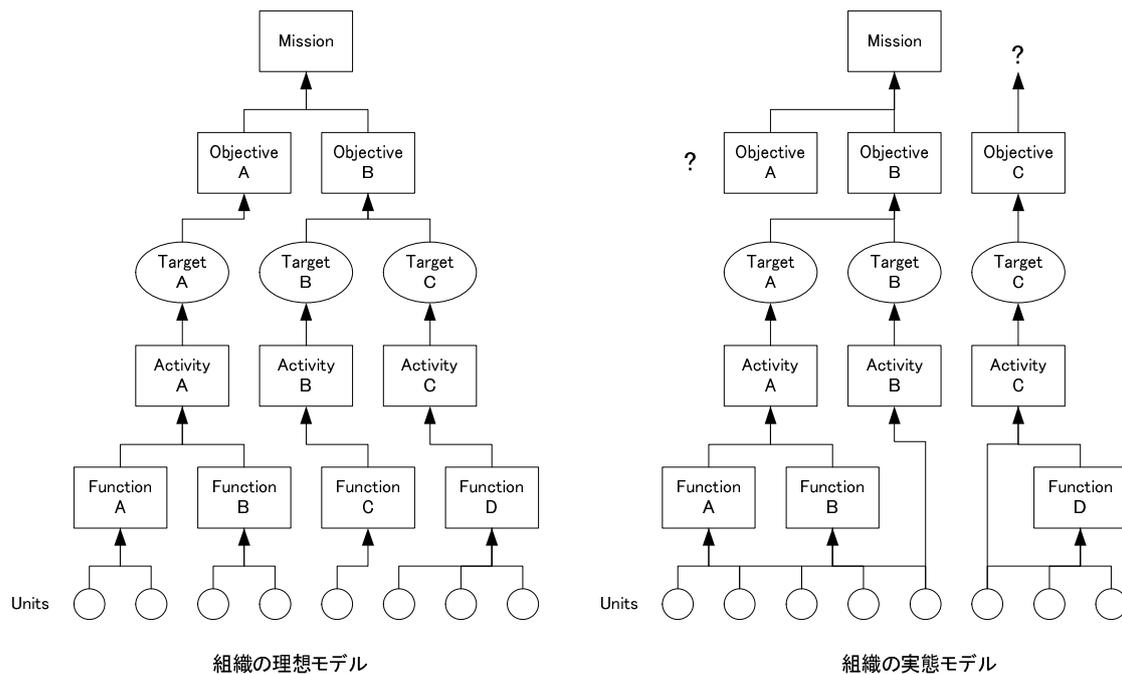


図 3-44 組織の理想モデルと実態モデル

出典：財団法人政策科学研究所作成資料

上の図 3-44 は組織の目的に応じたさまざまな機能に着目して構築した組織の理想モデルと、それと対比した形で表した組織の実態モデルを表したものです。理想モデルでは、組織のミッション、目的 (objective) に応じて事業等の活動 (activity) が位置づけられ、そのための機能 (function) が明確にされ、各ユニットは担当する機能が明確になっています。この状態を「組織のプログラム化」と呼びます。

一方、組織の実態モデルでは、必ずしもミッションに対応した事業目的を持たず、事業そのものが自己目的化しています。また、目的を達成するための活動や機能が明確に位置づけられておらず、各ユニットも複数の業務を担当するような構造になっていて、各ユニットの業務に対する貢献度も分かりません。

こうしたモデルは組織の目的や活動、機能をどのような事業レベルで定義するかによって描かれるパスが異なりますが、理想モデルと実態モデルとを分ける最大の特徴は、上位目的や上位の活動に対して手段となる活動やユニットが独立な要素として機能しているかどうかという点です。実態モデルでは、事業目的やユニットの業務に重複があり、効率的な運営がなされていないと考えられます。

組織の理想モデルと実態モデルはどのような組織においても通常は乖離しており、また、接近している場合にも、常に理想モデルの再構築と、実態モデルをその理想モデルに近づけていくというシステム・ダイナミクスが働いています。組織の理想モデルに実態モデルを近づける行為を本来「組織経営」と呼びます。

中長期計画・年度計画で描かれる組織像や事業のイメージは「理想モデル」でなければなりません。現実の組織が上に挙げたような「実態モデル」であったとしても、理想モデルに近づける努力を行う必要があります。

### (3) 機関のアウトカム・インパクトとは？

機関評価との対応で考えれば、機関のアウトカムやインパクトとは何かという問題を考慮する必要があります。先のミッションから活動の連鎖に至るモデルと、評価の観点から整理した活動からアウトカムやインパクトに至る連鎖とを比較してみましょう。

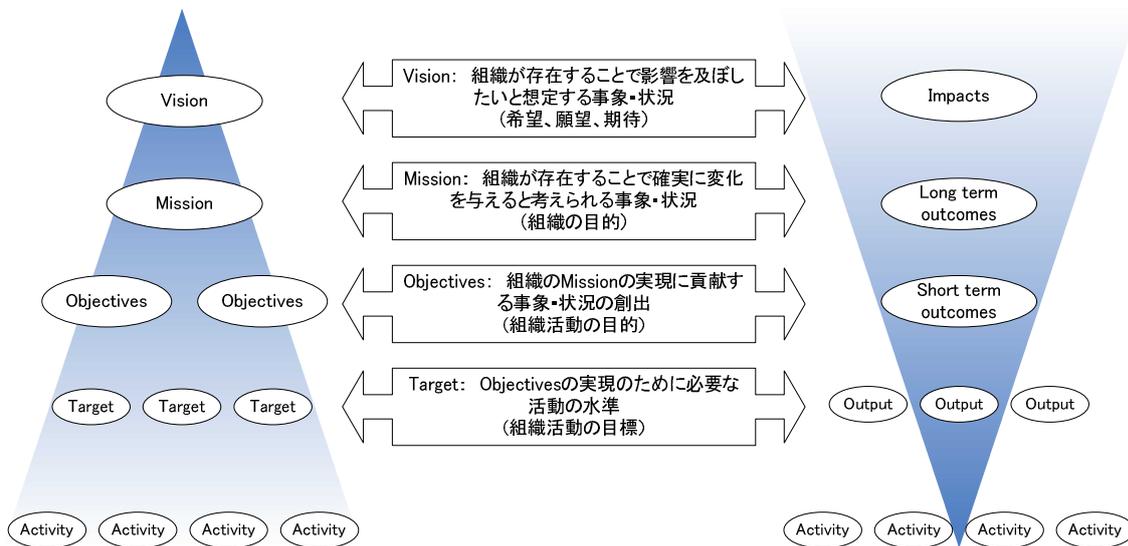


図 3-45 組織のプログラム化と機関評価の観点

出典:財団法人政策科学研究所作成資料

図 3-45 は組織のミッションから個別の活動までに至るプログラム化の方法（プログラム・デザイン）と、組織の個別の活動から影響経路を追っていく機関評価の観点とを対比して表したものです。

図 3-45 の左側から見ていきましょう。ビジョン (vision) とはプログラムや組織が究極的に達成したい事象、あるいは状況の変化を意味し、組織の意思 (will) を表明したものです。ただし、その組織が存在しなければ実現しないという類のものではなく、ビジョンとして描かれる事象や状況は他の主体やプログラムの影響を受けて成立します。したがって、中長期目標・中長期計画等における表現としては、「～に資する」、「～に貢献する」、「～を目指す」といった記述の仕方になります。ビジョンとはその実現のための行動を伴うものであり、行動を伴わないものは「祈り (prayer)」でしかありません。ビジョンが達成された状態を通常、「目的の地 (Goal)」と呼びます。

ミッション (mission) とは、そのプログラムや組織が存在することで確実に影響を及ぼすと言うことのできる事象、あるいは状況の変化そのものを意味します。したがって、目的表現としては、「～を果たす」、「～を実現する」といった表現が用いられます。ミッションはプログラムや組織が目的として表明したことに対する結果責任を負うものであるため、この段階までが自己評価の責任の範囲内にあるものと考えられます。

目的 (objectives) はミッションの実現に貢献すると考えられる活動の結果のイメージです。組織で展開されるさまざまな活動（事業、業務）が何をもたらし、ミッションといかに結びついているのかが示されなければなりません。つまり、目的を上位ミッションに対する手段として位置づけることで、連鎖を繋げていきます。中期計画等では「〇〇を通じ、～を

果たす」といった表現の「〇〇」に相当する部分となります。

目標 (target) は目的実現のために達成しなければならない活動の水準を表しています。年度計画等において、「平成△△年までに〇〇の効率を□□%達成する」という表現で客観的に示されます。

プログラム・デザインは、目標を達成する手段である具体的な活動 (activity) を位置づけることまでが最小限の作業範囲となります。組織の場合であれば活動は事業や業務に相当します。事業や業務に投入されるリソースがどれくらいか、また、どのような体制・マネジメントで実施すべきかといった部分を設計することは通常、プロジェクト・デザインと呼ばれます。プログラム・デザインとプロジェクト・デザインが結びついて初めて、「サービスの提供や業務の質の向上」と「業務運営の効率化」とを同時に議論できます。

個別の業務に関して「業務運営の効率化」が把握できれば、中長期目標の検討事項である「4) 財務内容の改善に関する事項」がリアルな数値の積み上げによって検討することが可能となります。

他方、ビジョン、ミッション、目的、目標が評価の観点からどのような対象として位置づけられるかを示しているのが図 3-45 の右側になります。個別の活動が生み出したものがアウトプット (outputs) であり、これは目標と対応する形で評価されます。また、目的と対応するのが短期的なアウトカム (short-term outcomes) になります。短期的なアウトカムの把握までは事業レベルの評価 (プログラム評価等) を活用することで得られます。

ミッションと対応するのが長期的なアウトカム (long-term outcomes) です。これは、機関が存在しなかった場合を非実施仮説にして、当該機関のサービスの提供がもたらした効果を計測する試みになります。研究開発機関であれば、実施した研究開発の成果の総体 (論文や特許) が科学技術領域においてどの程度のシェアを占め、同領域の発展に貢献したかを精査することで検証可能でしょう。資金配分機関であれば、研究開発助成や補助金の受け手 (カスタマー) に対するアンケート等で、公的研究資金が果たした役割について評価してもらうなどの方法が考えられます。

ビジョンに対応するのはインパクトです。インパクトのレベルでは、我が国の科学技術水準に関する国際比較など、マクロ的な視点で評価することになりますが、当該機関のみの貢献に帰することはできません。機関評価においてインパクトまでの立証責任はなく、内外の科学技術政策に関する研究成果に委ねられることになるでしょう。むしろ、そのような研究成果を引用し、機関が提供するサービスの影響が結果としてどのように現れているかどうかを確認するにとどまると考えられます。

#### (4) 活動と組織構造

組織のプログラム化と評価の観点との関係を踏まえた上で、機関評価のために必要な措置は、組織の活動と組織構造に関する対応関係の把握です。

組織の理想モデルと実態モデルの説明で触れたように、現実の組織の活動はミッションから目的、目標、活動へと独立な関係性を保って展開されているわけではありません。まずは組織の中の個別のセクションがどのような業務を担当しているのかについて、業務分析を行う必要があります。

図 3-46 は、古い情報になりますが、独立行政法人製品評価技術基盤機構 (NITE) の業務と担当部署との対応関係を描いたものです。

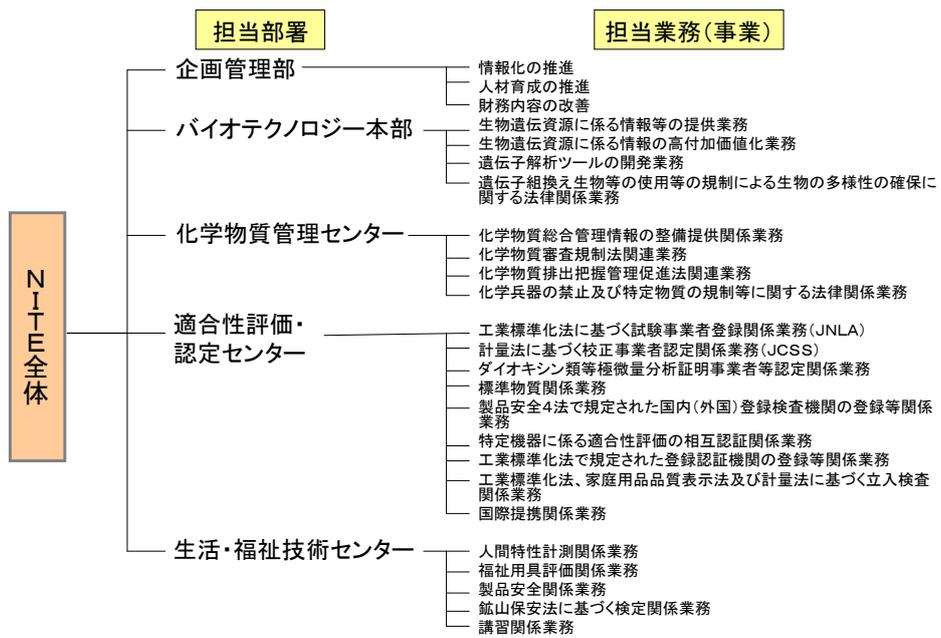


図 3-46 担当部署と担当業務の把握：製品評価技術基盤機構（NITE）の例

出典：平澤冷(2007),「独法評価への対応～製品評価技術基盤機構での経験から～」, 日本原子力研究開発機構講演資料

しかし、これらは主務とする業務を描いたものであり、それぞれの業務に何かしらのかたちで関与するセクションを業務別に展開すると、各セクション別の業務が 660 項目にも上りました (図 3-47)。

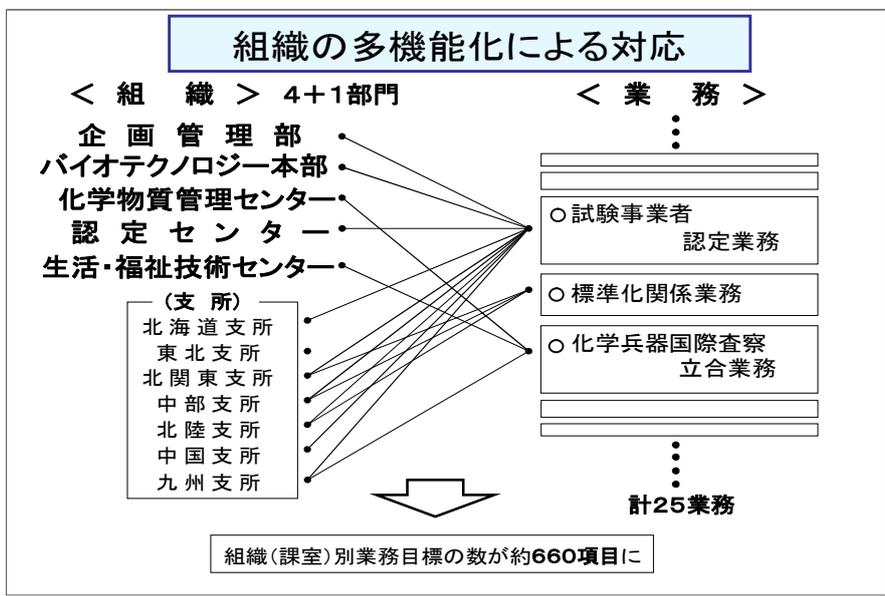


図 3-47 NITE の業務分析の結果

出典：平澤冷(2007),「独法評価への対応～製品評価技術基盤機構での経験から～」, 日本原子力研究開発機構講演資料

NITE では業務項目別に業務目標を定め、各セクションの業務に関わる費用を推計し、業務目標と対応させることで業務効率を計測するためのデータベースを 2 年間掛けて作成しました。

さらに、NITE では組織活動のプログラム化を実施しました。NITE の提供するサービスについて、アウトカムの観点から事業分野を類型化し、事業分野別のアウトカムとコストデータを対比することで業務効率の改善効果を数値に基づいて客観的に示しています。これらの改善効果を組織全体で積み上げることで、NITE 全体のコスト改善効果を示しています(図 3-48 および図 3-49)。

NITE の機関評価の経験<sup>24</sup>は現時点における独立行政法人評価の理想的な事例と言えます。しっかりとしたシステム設計の下、組織の業務分析、活動基準会計手法(ABC)による業務費用の推計、国民の視点(アウトカム)からの事業分野の類型化とミッションの再定義、機関全体のパフォーマンスの計測を行うことができたのです。

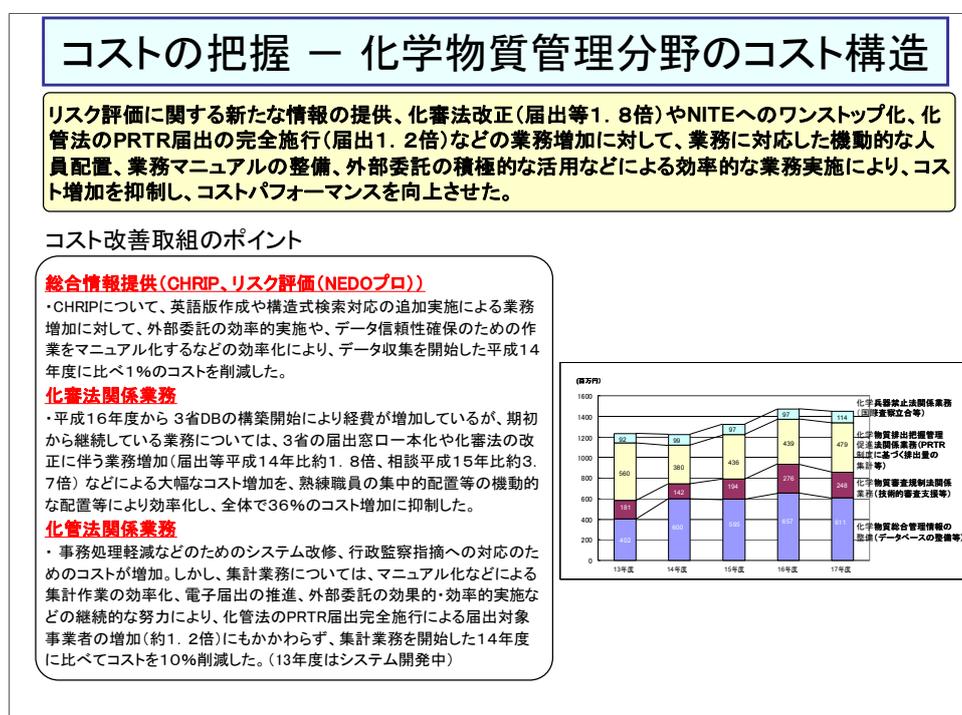


図 3-48 NITE における事業分野別のアウトカムとコストデータの対比

出典: 平澤冷(2007), 「独法評価への対応～製品評価技術基盤機構での経験から～」, 日本原子力研究開発機構講演資料

<sup>24</sup> <http://www.nite.go.jp/hyouka.html>

## コストの把握 — nite全体のコスト構造

- 期初からの継続業務については、5年間トータルで9.2%のコスト削減を達成。
- 第一期中に開始したBRC業務では中期目標で想定していなかった海外微生物収集事業の展開など当初の想定外の大きな成果を上げ、MLAP認定業務、講習関係業務では短期間での業務立ち上げに加え業務量の大幅な変動への着実な対応を図ることにより、高いパフォーマンスを示すとともにコストの増加抑制を図った。
- 期初からの業務においても、ゲノム解析業務での独自の前処理方法の開発や独自ツールの開発による効率化の追求、化審法業務での三省共通事務局化による業務量増加への効率的対応、認定業務での複数のプログラムでの横断的な内部職員の活用による業務量変動の吸収、製品安全業務での事故情報収集・検証件数の増加という環境下での社会的インパクトの高い情報発信とコスト削減の両立など、高いパフォーマンスの提供とコスト削減を両立した。

### コスト改善取組のポイント

#### ○バイオテクノロジー分野

- ・収集菌株の迅速な簡易同定方法の確立
- ・独自の前処理方法の開発や独自ツールの開発によるゲノム解析業務での効率化の追求

#### ○化学物質管理分野

- ・CHRRIP構築でのデータ信頼性確保のための作業のマニュアル化による効率化
- ・化審法業務での三省共通事務局化による業務量増加に対し、職員の機動的配置等による効率的対応

#### ○認定業務

- ・複数のプログラム間での横断的な内部職員の活用による業務量変動の吸収と外部審査員の育成・活用の推進

#### ○人間生活福祉分野

- ・製品安全業務において事故情報収集・検証件数が大幅に増加する中で「誤使用ハンドブック」等の社会的インパクトの高い情報発信とコスト削減を両立
- ・受講者変動の著しい講習業務において増加コスト抑制とサービス向上の両立を達成。

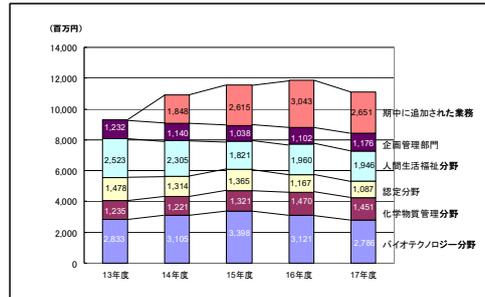


図 3-49 NITE 全体のコスト改善効果

出典：平澤冷(2007),「独法評価への対応～製品評価技術基盤機構での経験から～」,日本原子力研究開発機構講演資料

### 3.8.4 評価の方法

#### (1) 業務分析の実施

##### 1) 業務対応表の作成

機関評価のためには、機関で実施している業務を全て洗い出します。そして、業務を担当している部署(セクション)を業務毎にリストアップします。この作業では、次のような業務対応表をデータベース形式で作成することが望ましいでしょう。

表 3-20 業務対応表の例

NO.	事業分野名	事業名	業務名	業務内容	担当セクション	所属事業部
T-00001	○○推進分野	△△事業	××業務	....	□□推進室	○○事業本部
T-00002	○○推進分野	△△事業	××業務	....	◇◇事業所	○○事業本部
T-00003	○○推進分野	△△事業	●●業務	....	■■推進室	○○事業本部
T-00004	○○推進分野	△△事業	●●業務	....	■■推進室	○○事業本部
T-00005	○○推進分野	△△事業	●●業務	....	■■推進室	○○事業本部
T-00006	○○推進分野	△△事業	●●業務	....	■■推進室	○○事業本部
T-00007	○○推進分野	●●事業	●●業務	....	■■推進室	○○事業本部
T-00008	○○推進分野	●●事業	●●業務	....	■■推進室	○○事業本部
T-00009	○○推進分野	●●事業	●●業務	....	■■推進室	○○事業本部
T-00010	○○推進分野	●●事業	●●業務	....	■■推進室	○○事業本部
T-00011	●●推進分野	●●事業	●●業務	....	■■推進室	○○事業本部
T-00012	●●推進分野	●●事業	●●業務	....	■■推進室	○○事業本部
T-00013	●●推進分野	●●事業	●●業務	....	■■推進室	○○事業本部
T-00014	●●推進分野	●●事業	●●業務	....	■■推進室	○○事業本部

出典：財団法人政策科学研究所作成資料

業務対応表はボトムアップで作成すると良いでしょう。業務を統括する最小単位のセクション毎に、年間を通じて発生する業務をすべてリストアップしてもらいます。次に、その業務がどのような事業に関連したものかについて対応付けします。

さらに具体的な事業名が、機関においてどのような事業分野に位置づけられるかについて対応付けします。この時、事業分野に関して、ミッションを実現するための手段として位置づける（アウトカムの視点から位置づける）ことが出来れば、組織のプログラム化に役立ちます。

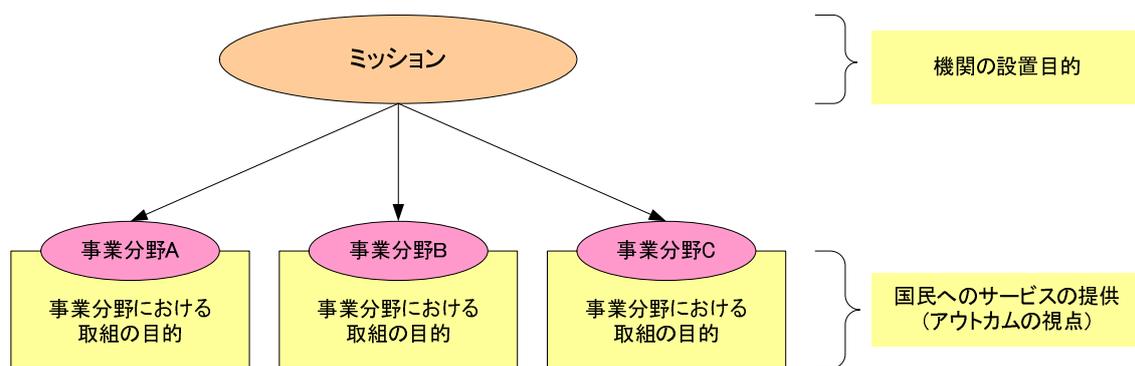


図 3-50 事業分野の再設定と組織のプログラム化

出典：財団法人政策科学研究所作成資料

業務対応表を各フィールド別に集計した場合、次のような対応関係の把握が可能になります。

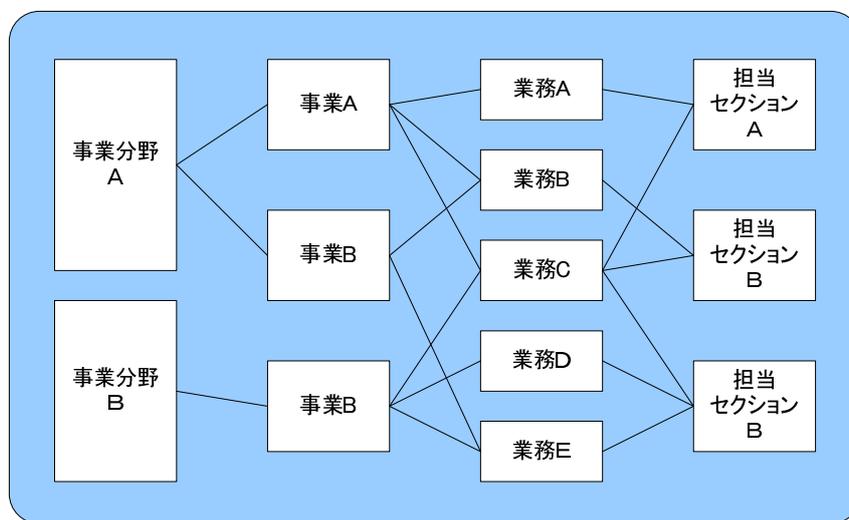


図 3-51 業務対応表の集計結果

出典：財団法人政策科学研究所作成資料

担当セクション別に集計した場合、担当セクションから実施している業務の一対多関係が導かれるでしょう。逆に業務別に集計した場合には担当セクションへの多対多関係が見えてきます。業務から事業への対応も多対多関係であることが分かります。

業務分析では、組織の活動の基本単位を業務に求めた場合、事業に対してどのような重複

や担当セクションの関わりがあるかという部分が組織の実態モデルとして見えてきます。この関係性を組織の理想モデルに近づけるためには2つの方法が考えられます。

- ① 事業から業務への関係がなるべく一対多関係になるように、業務の重複を避ける。
- ② 担当セクションに対して主たる業務を割り当てる。

## 2) データベースの関連づけ

業務対応表にすべての業務を記入することができたら、インプットデータの登録と、アウトプットデータの登録を行います。インプットデータの最小単位はおそらくどのように優秀な経理システムを導入していたとしても、事業部単位の把握が限界でしょう。

表 3-21 インプットデータの登録

NO.	コスト	万円	所属事業部
C-00001	人件費	2,000	〇〇事業本部
C-00002	設備費	150	〇〇事業本部
C-00003	⋮	⋮	⋮
C-00004	⋮	⋮	⋮
C-00005	⋮	⋮	⋮
C-00006	⋮	⋮	⋮

出典：財団法人政策科学研究所作成資料

また、アウトプットデータの最小単位は実施事業単位であることが想定されます。これらのデータは研究開発評価のデータをそのまま流用できます。

表 3-22 アウトプットデータの登録

NO.	事業名	アウトプット分類	件数
O-0001	△△事業	特許申請数	10
O-0002	△△事業	論文	12
O-0003	△△事業	試作機	3
O-0004	△△事業	⋮	⋮
O-0005	⋮	⋮	⋮
O-0006	⋮	⋮	⋮

出典：財団法人政策科学研究所作成資料

これらのデータを先の業務対応表の集計結果と接続すると、下のようなりレーショナル・データベース（RDB）を作成することができます。

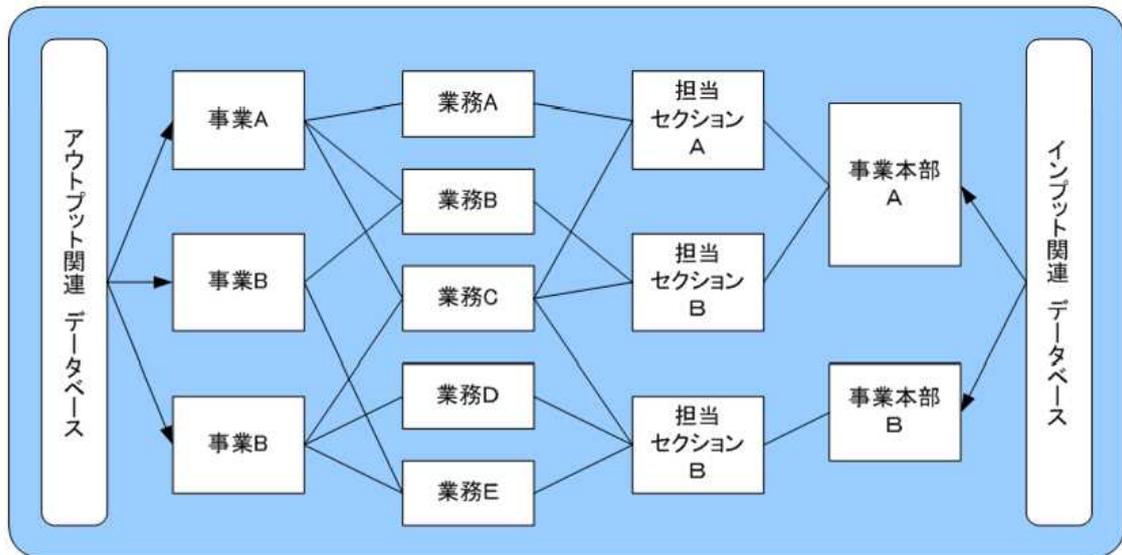


図 3-52 業務分析のためのデータベースの作成

出典:財団法人政策科学研究所作成資料

上のような RDB を作成して、ようやく、アウトプットとインプットの対応を、業務という機能を媒介にして関係づけることが可能になり、「業務運営の効率化」を評価することができます。

## (2) 活動へのコストの割り当て

業務分析を実施した結果、インプットとアウトプットの対応関係については見えてきましたが、ある業務ないしある事業に一体いくら年間に費やされたかを明らかにするためには管理会計手法を適用する必要があります。代表的な管理会計手法が活動基準会計（ABC: Active Based Counting）と呼ばれるものです。

### 1) 活動基準会計（ABC）

活動基準会計とは、高度に発達した組織活動において管理費用や調整費用などの間接経費が膨張するに従い、無視できないコスト要因になってきたことを経営上の問題として捉える必要性から生み出された原価計算手法のことです。活動基準会計によって、活動ごとに把握された原価情報などを活用し、コストの視点から活動の管理に重点を置くことが可能になります。活動の分析を通じてプロセスのムダ（非付加価値的活動）が明らかにされるので、組織の実態モデルをリエンジニアリングする際に役立つ手法として注目されています。

活動基準会計の手続は次の通りです。最初に、業務の原価として考慮しなければならないコストを列挙します。コストデータは会計年度において確定された財務データから抽出します。

次に、コストデータを部門別に配分します。積み上げでデータを蓄積している場合には問題ありませんが、コストデータによっては把握できない場合は、部門全体の年間の予算比率等によって配分します。同様に、担当セクション別のコスト比率も推計します。

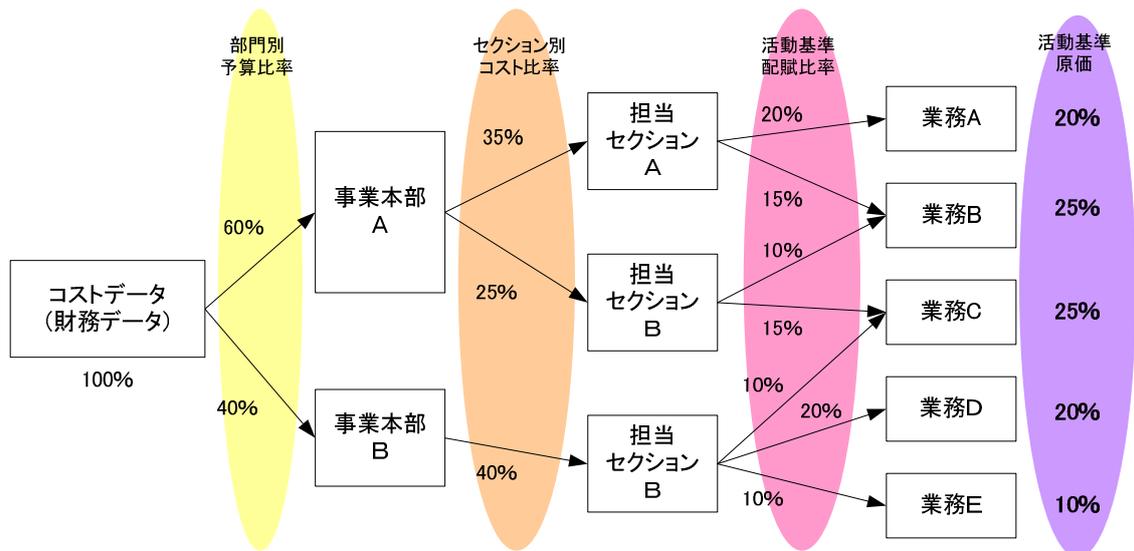


図 3-53 活動基準会計の考え方

出典: 財団法人政策科学研究所作成資料

担当セクションが業務に対して費やしている費用については、何かしらの配賦基準（コストドライバー）を基に配分する必要があります。活動基準原価計算では、この配賦基準を活動の度合いをもっとも良く表している指標を設定することで代理変数とします。例えば、労働時間（人・時）などが代表的な例です。

すべての業務についてコストが配分されたら、業務別にコストを集計します。これが、活動基準原価と呼ばれるものになります。

業務分析によって、機関の活動は「業務」－「事業」－「事業分野」への対応が図られています。業務の活動基準原価を積み上げることで、さまざまな観点からコストの集計が可能になり、アウトプットやアウトカムとの対比が可能になります。

### (3) マネジメントの評価

業務運営の効率化については、実施している業務それ自体の目標水準が定められていなければなりません。業務分析において業務をリストアップすると説明しましたが、それを業務別に推計するという作業は、実は業務の戦略的な類型化が前提となっています。

業務の戦略的な類型化とは、マネジメントの観点から業務がどのように位置づけられるかという問題を解くことです。こうした観点に有効な手法として、バランススコアカード（BSC）によるマネジメント評価があります。

バランススコアカードについては、本テキストの 3.4.1 において詳細に解説していますので、ここでは重複して解説することを避けたいと思います。

### (4) スタッフ組織の業務の位置づけ

業務分析で「事業」－「業務」－「担当セクション」の対応が比較的可能なのはライン部門であると言えます。ところが、研究開発機関や資金配分機関では、企画調整部門、研究開発評価部門、戦略策定部門等、多くの事業に関わるスタッフ組織が重要な役割を果たしてい

ます。

これらの部門に業務分析を単純に適用すると必要以上に多くの業務単位を作り出してしまふこととなります。このような問題を解決するために、担当セクションではなく、実際の業務を遂行するための機能を重視した組織構造を仮想的に設計し、業務分析に組み込む方法が考えられます。

#### 1) マトリックス型組織とクロス・ファンクショナル・チーム

図 3-54 はフィンランド科学技術庁 (TEKES) の組織構造を図示したものです。TEKES は日本でいうところの NEDO のような資金配分機関です。2006 年時点のものになりますが、TEKES では典型的なマトリックス型組織を構築しています。

図の縦軸は資金配分業務 (Core Processes) が採られ、横軸は事業領域 (Technology and Research Area) が採られています。また、右下には戦略形成・分析部門が両方の業務の軸にコミットするような形で組織構造を示しています。

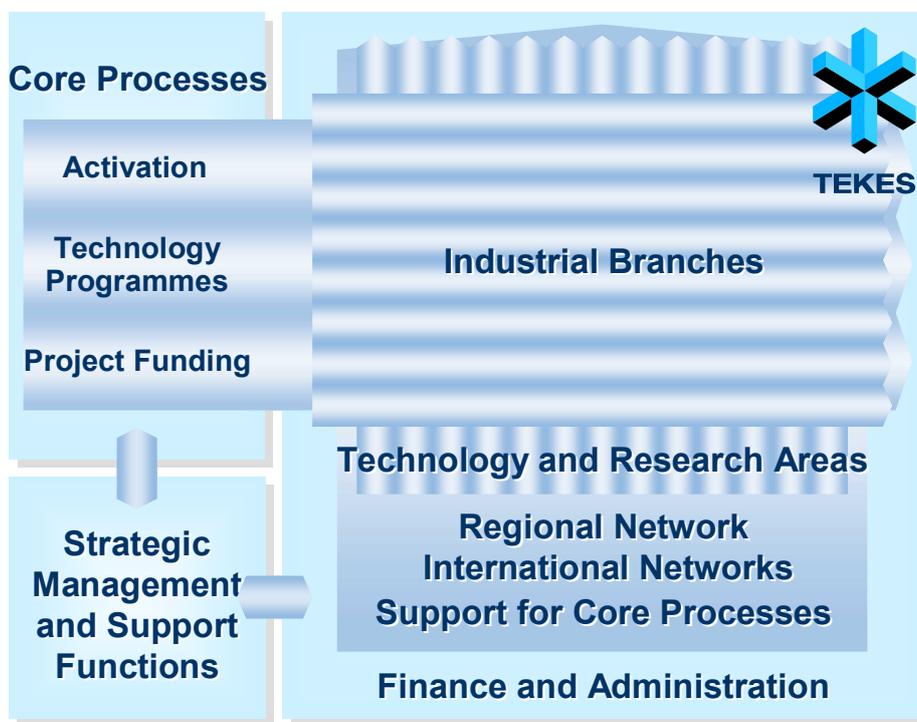


図 3-54 マトリックス型組織 (TEKES の例)

出典: TEKES ヒアリング資料(2006.1)より抜粋

マトリックス型組織は、形式的には他の組織と同様に各セクションが独立したツリー構造に位置づけられています。しかし、機能的にはスタッフ組織とライン組織が区別され、両方の組織が事業を共有するかたちで業務を遂行するようになっています。

このとき、組織の垣根を越えるようにそれぞれのセクションからプロジェクトに応じてメンバーを募り、チームを作るような実施体制をクロス・ファンクショナル・チーム (CFT) と呼びます。CFT は日産のゴーン社長がフェニックス計画で導入したマネジメント手法として一躍有名になりました。

CFT の考え方を業務分析に導入すると、次のような関係性の把握ができるようになります。

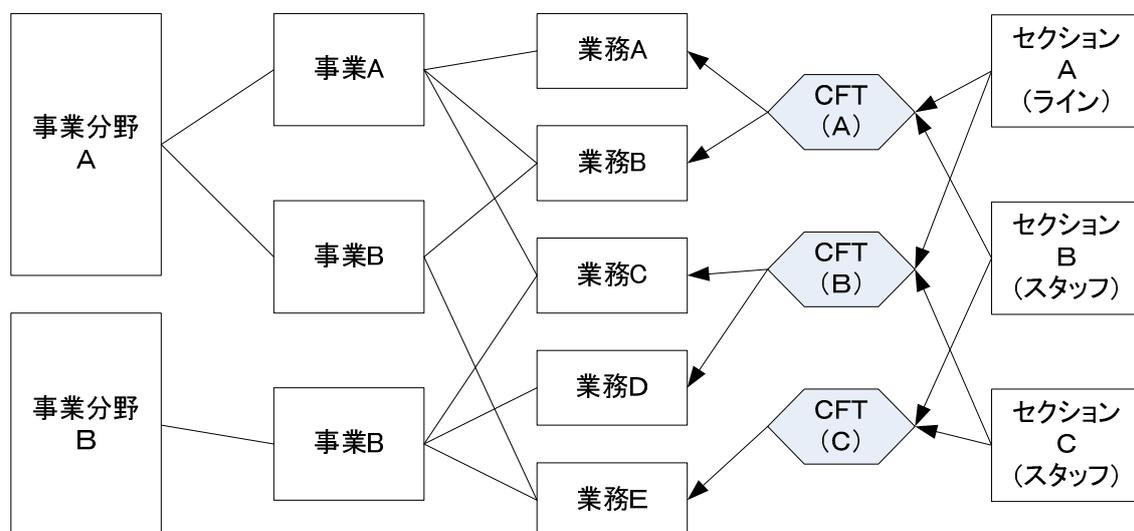


図 3-55 マトリック型組織と業務分析 (例)

出典:財団法人政策科学研究所作成資料

図 3-55 では、業務を直接担当のセクションに関連づけるのではなく、セクションのメンバーから構成される CFT に関連づけています。このようにすることで、CFT が実施すべき業務をセクションに割り当てるのが間接的に可能になり、かつ、CFT が媒介となることで割り当ての組み合わせ数を大幅に減少させることができます。つまり、その後の活動基準原価計算の手間が掛からなくなります。

CFT は必要に応じて結成されるため、組織がフラットでかつ、機動的な構造になっていなければ機能しません。しかし、一度このような組織構造を作り上げることができれば、業務効率の評価の対象を CFT に絞り込むことができるため、機関評価の手続を大幅に短縮することができます。

組織構造にはこれが正解という理想形はありません。経営学のコンティンジェンシー理論では、「どのような組織にも幅広く適用できる最適な組織構造というものは存在せず、組織は様々な要因に応じて組織構造を考え、見直していかなければならない」とされています。CFT の考え方は、組織のプログラム化を一層進めたものであると言えますが、こうした取り組みを行っている機関と比較などを行い、適宜組織構造を見直していくことも、機関評価に求められる役割の 1 つであると言えるでしょう。