

総合科学技術会議
第31回評価専門調査会

研究開発評価の諸相と展開状況

— 今後の課題は何か —

中央合同庁舎 4号館
平成16年1月20日(火)

東京大学名誉教授
平澤 冷
rh@rhirasawa.info

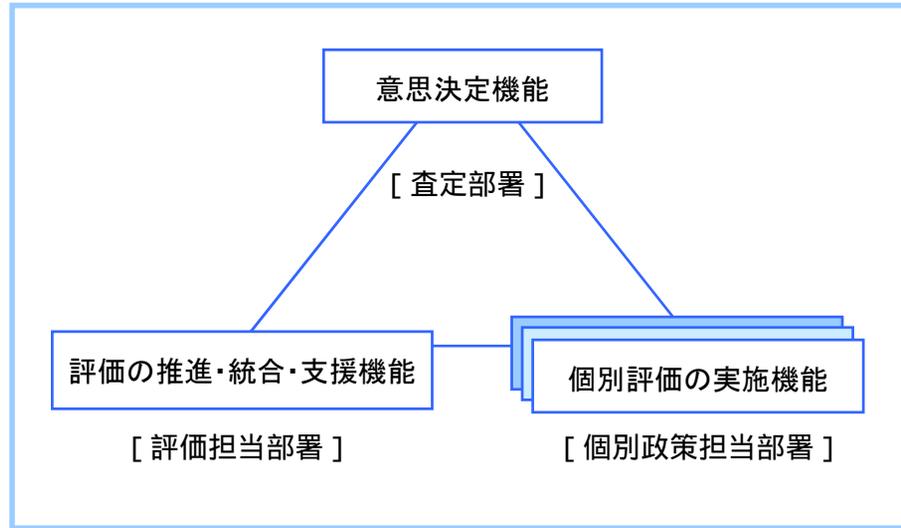
我が国における公的研究開発評価制度の展開

- 65 農林省研究所レビュー
- 66 通産省大型プロジェクト評価
- 71 科学技術会議5号答申(ソフトサイエンスの振興)
- 84 科学技術会議11号答申:研究評価の充実
 - 85 臨時行政改革推進審議会答申(科学技術政策大綱提起)
- 85 科学技術会議政策委員会・研究評価指針策定委員会の設置
- 86 「研究評価に関する基本的考え方」、「研究評価のための指針」(研究評価技術策定委員会)
- 87 科学技術会議政策委員会:大規模プロジェクト評価の検討の進め方
- 87 科学技術会議13号答申:国研問題
- 89、90 厚生省評価マニュアル
 - 92 科学技術会議19号答申(ソフト系科学技術の研究開発基本計画)

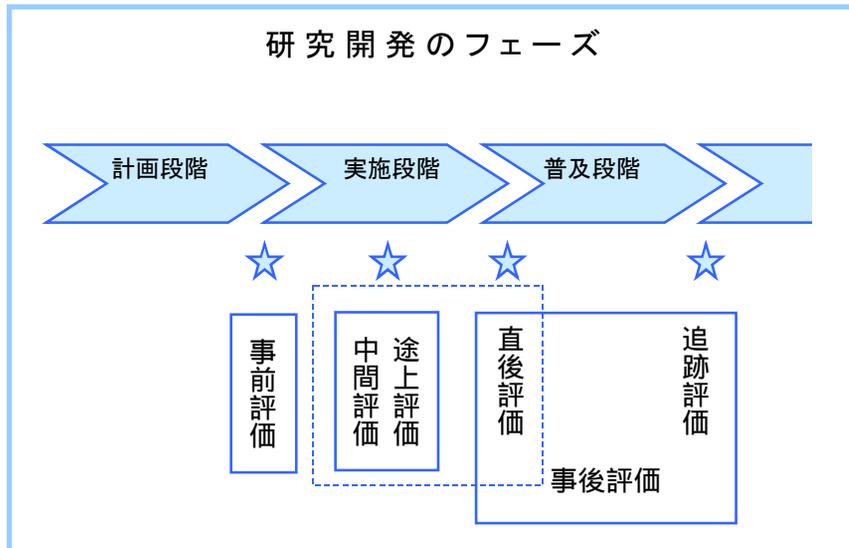
- 95 科学技術基本法公布・施行
- 96 (第1期)科学技術基本計画(閣議決定)
- 97 国の研究開発全般に共通する評価の実施方法の在り方についての大綱的指針(内閣総理大臣決定)
- 97 郵政省・農水省・通産省・科学技術庁・文部省などの評価指針
- 98 厚生省・環境庁・運輸省・防衛庁など評価の指針
 - 98 中央省庁等改革基本法(施行)
- 98 科学技術会議による評価実施状況のフォローアップ(以降毎年実施)
 - 99 各省庁設置法等からなる中央省庁改革関連法成立
 - 99 独立行政法人通則法(公布)、01.04(施行)
 - 99 経済産業省における政策立案・評価指針など

- 00.04 大学評価・学位授与機構設置法(施行)
- 01.01 新府省体制発足
- 01.03 (第2期)科学技術基本計画(閣議決定)
 - 01.01 内閣府・総合科学技術会議設置法(施行)
 - 01.04 独立行政法人発足
 - 01.06 行政機関が行う政策の評価に関する法律(政策評価法)成立(02.04施行)
- 01.11 国の研究開発評価に関する大綱的指針(内閣総理大臣決定)
- 02.03 総務省情報通信研究評価実施指針
 - 01.03 文部科学省政策評価基本計画
- 02.03 防衛庁研究開発評価指針
 - 02.04 経済産業省政策評価基本計画
- 02.04 経済産業省技術評価指針
- 02.04 総合科学技術会議評価専門調査会による重要研究開発課題の評価
- 02.05 文部科学省における研究及び開発に関する評価指針

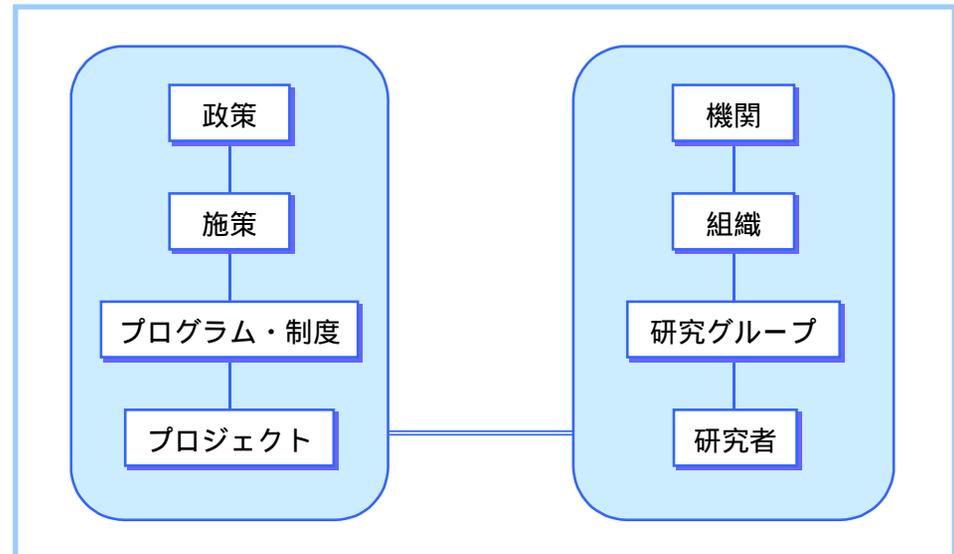
行政組織内部の評価関連機能



研究開発のフェーズと評価実施時期の関係



研究開発評価対象の階層構造



研究開発のカテゴリーと評価[事前]

- ・科学技術の枠内に目的を設定(シーズ目的・シーズアプローチ)
基礎研究・学術研究
知的フロンティアへの挑戦
- ・ニーズを見据えシーズの側からアプローチ(ニーズ目的・シーズアプローチ)
長期研究, 計画的・戦略的展開
シナリオ, ロードマップ, アクションプラン
「経済代替性」評価: 代替案の経済性予測
「費用対効果」分析: 投入経費と予想される効果の関係を分析
「ネットワーク」分析, 「システム分析」: 開発体制の連携形態・強度等の分析
「フェーズ管理」: シーズ・プッシュからニーズ・プルへの転換メカニズムの設計と整備
- ・ニーズからの発想(ニーズ目的・ニーズアプローチ)
短期研究, 動的展開
「コスト分析」: 原価企画等
- ・科学技術の基盤整備
基盤研究
ポテンシャルや体制の評価

科学技術関連政策のカテゴリーと評価対象

科学研究
技術開発

RTD Policy

科学と技術の
ポテンシャルとシーズ

イノベーション

Innovation Policy

	シーズとニーズを繋ぐ 広い意味の連携政策	連携推進のための 基盤整備政策	実用化推進 政策
市場内部性課題			
市場外部性課題 (含市場の失敗)			

研究開発評価の内部構造

評価の理念

支援的
内在接触
明示性

評価対象

政策
プログラム
プロジェクト

機関
組織
人材

能力	マネジメント	政策と戦略	
使命, 戦略, 目標 フォーサイト調査 予測	パフォーマンス モニタリング 達成度	アウトプット アウトカム 実績	アウトカム インパクト 見なおし
事前評価	中間評価	事後(直後/追跡)評価	

科学技術的価値 : 知見・知識の質
 経済的価値 : 投資回収率, スピルオーバー等
 社会的価値 : 問題解決

データと方法論

順序
尺度
評点化
経験的データ
既存統計
ケース分析
文献分析

計量文献分析
計量技術分析
計量経済分析
計量社会分析
統計分析
構造化分析
システム分析

ヒアリング
インタビュー
アンケート
事例調査

ピア・レビュー/パネル法
エキスパート・パネル法
評点法
比率評価法
比較評価法
指標法
システム評価法
ロジック評価法
レビュー法

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, エキスパート・レビュアー(社会経済性レビュアー等)
 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

調査分析評価手法の体系

		調査法		分析法	手法区分	評価手法
		情報収集	データ収集			
調査分析評価手法	単純評価	インタビュー(関係者) ヒアリング(専門家)		ケース分析	定性的評価	ピアレビュー法 エキスパートパネル法
		既存統計 実在データ			半定量的評価	評点法 比率評価法
	複合評価	アンケート調査 事例調査 インタビュー(関係者) ヒアリング(専門家) 顧客調査 社会調査		文献分析 ケース分析 社会分析 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, サイエンスリンク分析, 技術仕様分析 計量経済分析(エコメトリクス) 費用指数分析, 費用便益分析, 生産関数分析, 財務指標分析, 消費者余剰分析 計量社会分析(ソシオメトリクス) CVM, 汚染指数分析, 健康・疾病指数分析, 雇用指数分析 統計分析 相関分析, 因子分析判別分析, クラスター分析 構造化分析 マトリクス法, 階層化法関連樹木法, 行列分析, マッピング システム分析 モデリング, プログラミングシミュレーション, ネットワーク分析	定性的評価	ピアレビュー法 エキスパートパネル法
		アンケート調査 事例調査 既存統計 実在データ 申請・管理データ データマイニング その他のマネジメント			半定量的評価	評点法 比率評価法
					定量的評価	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アプローチ (対照群・比較群アプローチ), 事前事後比較, 非実行 仮説, ランキング, ベンチマーク, ポートフォリオ 指標法 単一指標法, 複合指標法
					総合的評価	システム評価法 シナリオプランニング, ロードマップ, ディビジョンマネジメント評価 ロジック評価法 ファクト評価法, ケース評価法, レトロスペクティブ評価法 (歴史的遡及) レビュー法 デルファイ法, フォーサイト, フォーカスグループ

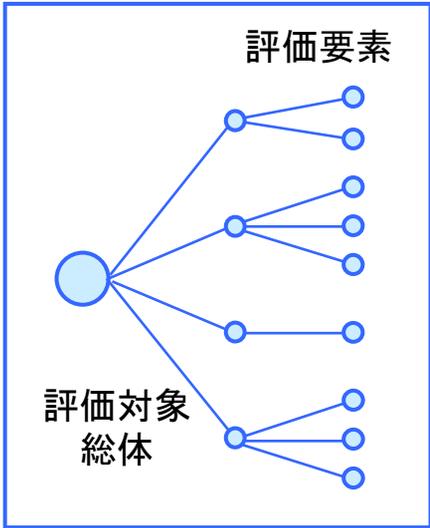
評価の定式化

評価体系

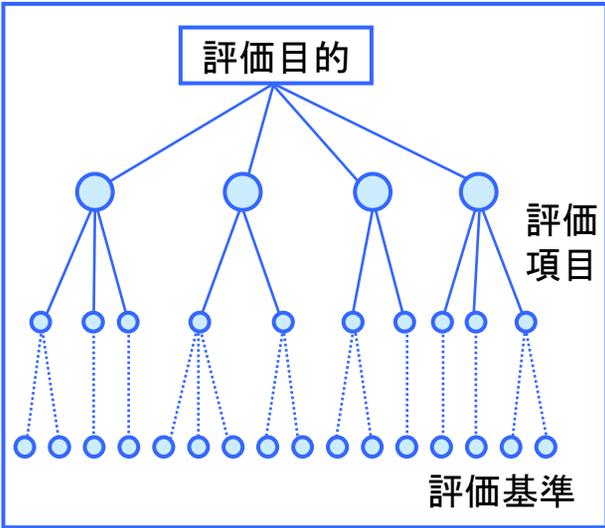
$$\begin{matrix} \text{評価対象系} \\ \left(\begin{matrix} \text{評価値} \end{matrix} \right) \end{matrix} \times \begin{matrix} \left(\begin{matrix} \text{重要度} \end{matrix} \right) \end{matrix} = \begin{matrix} \left(\begin{matrix} \text{評価結果} \end{matrix} \right) \end{matrix}$$

[個別評価] × [価値観] = [価値]

評価の定式



評価対象系



評価体系

評価項目と評価基準の総括表(1)

評価フェーズ の区分	評価項目	評価基準
事前評価	1. 目的の妥当性 <ul style="list-style-type: none"> a. 上位の目的に対する適合性 b. 上位の状況の変化に対する適合性 2. 目的の重要性 <ul style="list-style-type: none"> a. 内容の妥当性 3. 目的の実現可能性 <ul style="list-style-type: none"> a. 計画の妥当性 b. 手法の妥当性 c. 実施者の編成 d. 実施者の能力 e. 実施環境の適否 	<ul style="list-style-type: none"> ・ 上位の目的や計画に包摂されているか（予想される大状況の枠内に目的が設定されているか） ・ 上位の状況が変化している（変化が予想される）場合、それに対応した目的の設定になっているか ・ 期待される成果の意義やメリットは十分か <ul style="list-style-type: none"> - 科学的意義やメリット - 技術的意義やメリット - 経済的意義やメリット - 社会的意義やメリット ・ 成果を挙げるための手順は適切か ・ 成果の実現時期は適切か ・ 適用される方法論やアプローチは適切か ・ チームメンバーの構成は適切か ・ リーダーの指導力は十分か ・ チームメンバーの能力は十分か ・ 実施環境は十分支援的か
	4. 実施の可否	<ul style="list-style-type: none"> ・ 実施に移すべきか

評価項目と評価基準の総括表(2)

評価フェーズ の区分	評価項目	評価基準
途上評価・直後評価・追跡評価	1. 実績の状況 <ul style="list-style-type: none"> a. 成果の内容 b. 成果に対する寄与 c. コストの内容 2. 費用対効果 <ul style="list-style-type: none"> a. 費用対効果の様子 3. 成果の比較 <ul style="list-style-type: none"> a. 対計画（目標）比実績 b. 対基準年比実績 c. 対類似他者比 d. 対チャンピオンデータ比 4. コストの比較 5. マネジメントの妥当性 <ul style="list-style-type: none"> a. マネジメントの枠組みの妥当性 b. マネジメント手法の妥当性 c. リーダー等の指導力 d. メンバーの編成 e. メンバーの能力 	<ul style="list-style-type: none"> ・ 個別課題毎に、どのような成果が挙げたか。量的側面と質的側面について <ul style="list-style-type: none"> - アウトプット成果 - アウトカム成果 - インパクト成果 - 副次的成果 を挙げる ・ 個別成果毎の当該案件による寄与率はどの程度か ・ 個別成果毎に要した費用はどの程度か ・ 主要成果の費用対効果はどの程度か ・ 当該案件全体としての費用対効果はどの程度か ・ その他の比較（対計画比、対基準年比、対類似他者比等）における費用対効果の実績はどの程度か ・ 成果の達成度は計画（ないし目標）と比較した場合どの程度か ・ 基準年の成果に対する成果の成長率はどの程度か（対前年比等） ・ 類似した他の案件と比較し、成果はどの程度か（ベンチマーク、ランキング） ・ チャンピオンデータに対する成果の到達度はどの程度か ・ 3.の成果に対する比較と同種の比較 ・ マネジメントの枠組みやシステムは適切であったか ・ 取り組み方は適切であったか ・ リーダー等の指導性は十分発揮されたか ・ メンバーの構成は適切であったか ・ メンバーの能力は十分発揮されたか
	6. 継続の可否（途上・中間） 7. 目標や計画の修正（途上・中間） <ul style="list-style-type: none"> a. 目標の修正 b. 計画の修正 8. 目的の見直し（直後・追跡） <ul style="list-style-type: none"> a. 目的自体の見直し 	<ul style="list-style-type: none"> ・ 継続すべきか ・ 修正すべき目標は何か ・ 目標をどのように修正すべきか ・ 修正すべき計画は何か ・ 計画をどのように修正すべきか ・ 当初の目的を見直すべきか ・ 目的をどのように見直すべきか

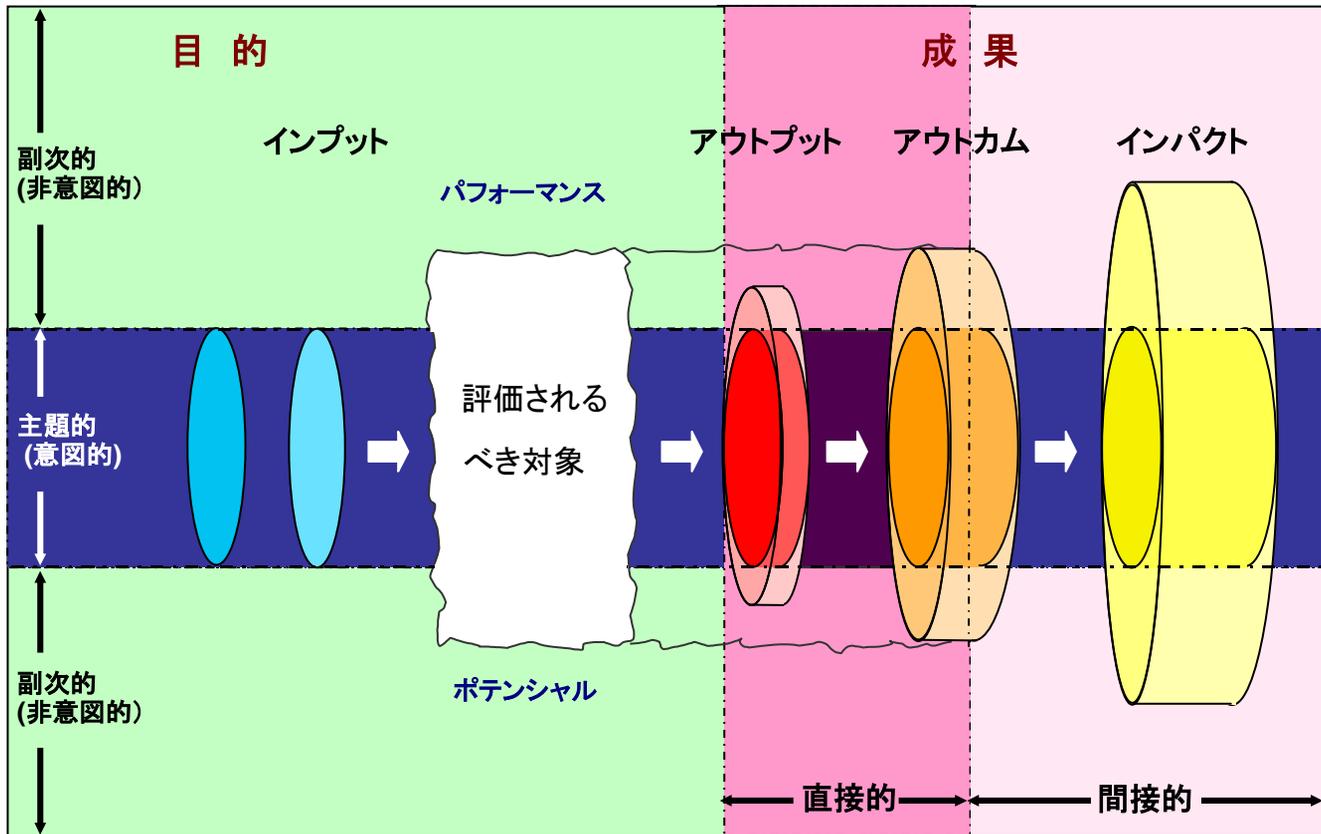
評価カテゴリー毎の主な評価目的と評価内容

	事前評価	(途中評価) 中間評価	事後評価	
			直後評価	追跡評価
従属プロジェクト	目的の適合性 [上位の枠組みとの照	目的や計画に対する到達度や達成度 [実績の評価, つまりパフォーマンス(研究の内容:「費用対効果」,「成果内容」,「コストの内容」とマネジメント(研究開発の運営のあり方)の評価] [対目的比, 対計画比の他に、対類似他者比, 対チャンピオンデータ比, 対過去比等の採用] [評価結果の計画や設定目標あるいは次期計画や上位の目標等の見直しへの反映]	通常「プログラム」の追跡評価の中で行われる	
プログラム 政策・施策 独立プロジェクト	合] 目的や目標自体の妥当性 [代替性等を含めた調査分析を前提とする]		実績の最終確認と設定された当初目標の妥当性 [成果全体(直接/間接, 目的内/目的外), つまりアウトプット, アウトカム, インパクト, 副次効果等の分析に基づく最終評価]	
最上位の政策	目的や目標自体の妥当性 決定プロセスの正当性 [フォーサイト等の社会に開かれた評価システム]		[目的や計画等の妥当性判断に基づく上位目標や次期計画等の見直し]	
機関	機関の新設は、「政策」や「施策」課題に類する ----- 機関内で発足させるプロジェクトの事前評価は「従属プロジェクト」の事前評価と同一		[事前, 中間, 直後評価のあり方に対する評価] 機関内で発足させたプロジェクトの追跡評価は「従属プロジェクト」の追跡評価に類する	

評価対象の原理的区分

	事前評価	途上（中間・直後）評価	追跡評価
政策	事前評価		
施策・制度 （プログラム）			
プロジェクト			
機関	年次計画	年次評価	年次評価
	期間計画	期間評価	期間評価
評価の視点	<ul style="list-style-type: none"> ・ 設定される枠組（目的）の 必要性 / 妥当性 有効性 / 効率性 期待される効果の予測 他のオプションとの比較 	<ul style="list-style-type: none"> ・ 計画の達成度（対計画比） 実績の把握 ・ 設定された枠（目的）内での 見直し 有効性と効率性 	<ul style="list-style-type: none"> ・ 設定された枠組（目的）自体の 見直し 有効性と効率性の他に、 必要性 / 妥当性 実績の把握の他に、 他のオプションとの比較 ・ 対応の改善

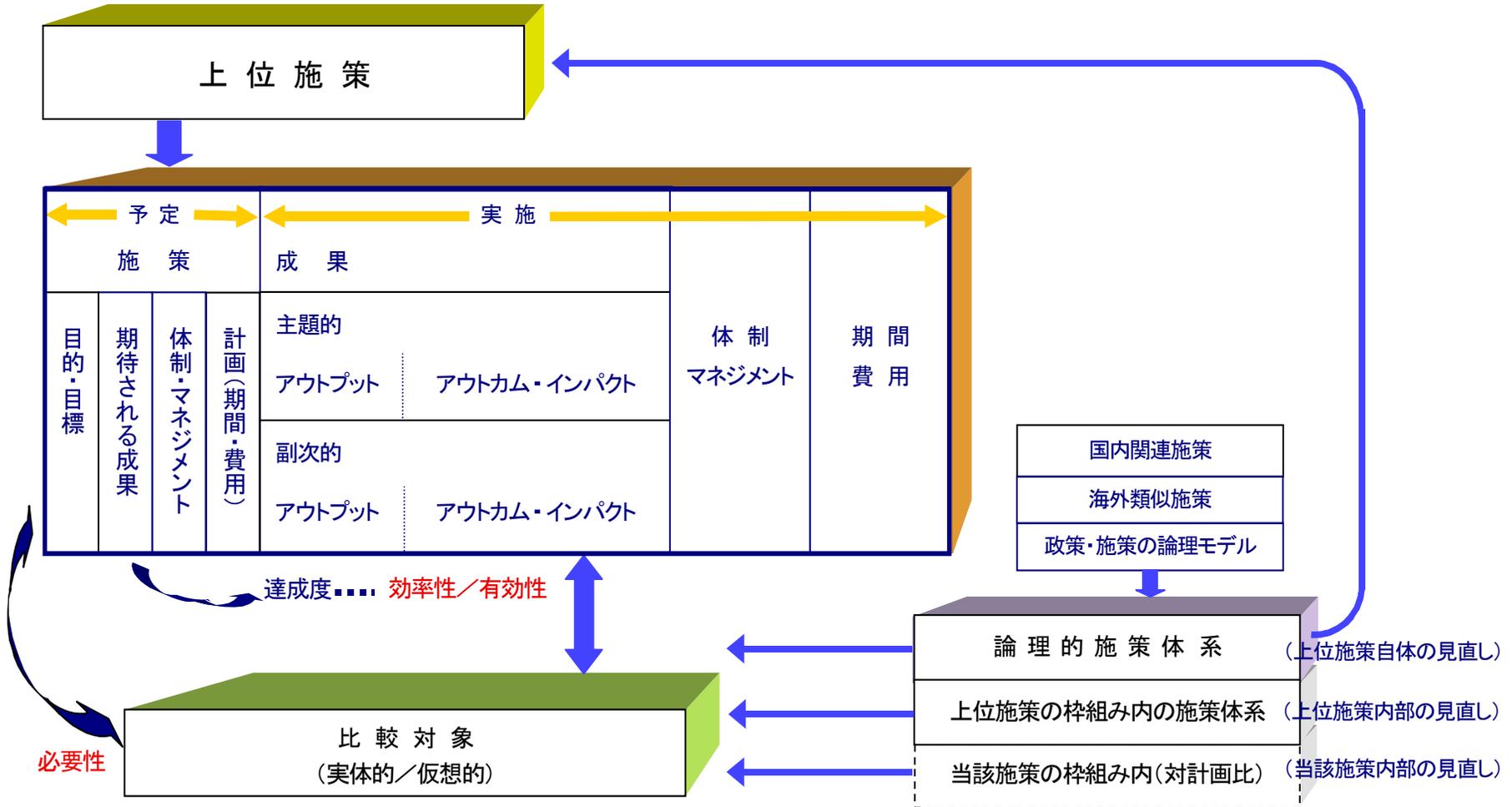
成果のカテゴリーと関連概念



途上(中間・直後)評価の評価項目と評価基準

評価項目			評価基準（計画に対する比較：達成度）	
実績	成果	内容	有効性	アウトプット / アウトカム / インパクト 主題的 / 副次的 質的 / 量的
		寄与		(当該案件の) 寄与率
	費用	(成果毎の) 費用 (その) 内部構造		
	体制	実施組織 実施者		編成 (の改善) ネットワーク (の目的適合的な拡大) 能力 (の増進)
	マネジメント		妥当性	枠組み / 手法
	コストパフォーマンス		効率性	経済効率性 (成果 / 費用) ミクロ経済性 (BETA) 生産性 費用対効果 (CBA)

施策追跡評価の基本スキーム



追跡評価の枠組み

評価項目	「実績」との比較対象	評価手法（例）	見直しの対象
目的達成度	計画 対基準年	単純比較法 単純比較法	体制 マネジメント
目的（事案）の有効性 / 効率性	コントロールグループ 類似他者 チャンピオンデータ	アディショナルリティ ベンチマーク、ランキング ランキング	目的 体制 マネジメント
目的（事案）の必要性 / 妥当性	非着手 代替案	シミュレーション システム評価法、シミュレーション	目的
上位の目的（事案）の妥当性		目的のメタ評価	上位目的

事前評価の枠組み(従属型プロジェクト)

評価項目	評価基準
目的の必要性 / 妥当性 上位の目的に対する適合性 上位の目的が想定する状況変化に対する適合性	上位の目的が計画に包摂されているか(上位の目的が予想する大状況の枠内に当該目的が設定されているか) 上位の目的の下にある関連施策との補完性(重複・欠落) 未展開施策に対する優先性は十分か 上位の状況が変化している(変化が予想される)場合、その状況に対応した目的設定になっているか
目的の重要性 内容の妥当性	期待される成果の意義やメリット(有効性/効率性)は十分か - 科学的意義やメリット - 技術的意義やメリット - 経済的意義やメリット - 社会的意義やメリット
目的の実現可能性 計画の妥当性	成果をあげるための手順は適切か(実現すべき目的に見合ったアプローチとなっているか) - 主要課題は、科学研究か、要素技術開発か、実用機能開発か、事業化か - 主要課題の克服に見合った計画となっているか 成果の実現時期は適切か

評価項目	評価基準
手段の妥当性	適用される方法論やアプローチは適切か
実施者の編成	チームメンバーの構成は適切か
実施者の能力	リーダーの指導力は十分か チームメンバーの能力は十分か
実施環境の適否	実施環境は十分支援的か
行政関与の必要性 公的資金の必要性	市場の外部性の程度 市場の失敗の程度 市場内部性課題の場合の費用対効果からみた必要性
公的対応の必要性	体制 マネジメント
実施時期の妥当性	緊急性、なぜ今か

事前評価の枠組み(独立型プロジェクト)

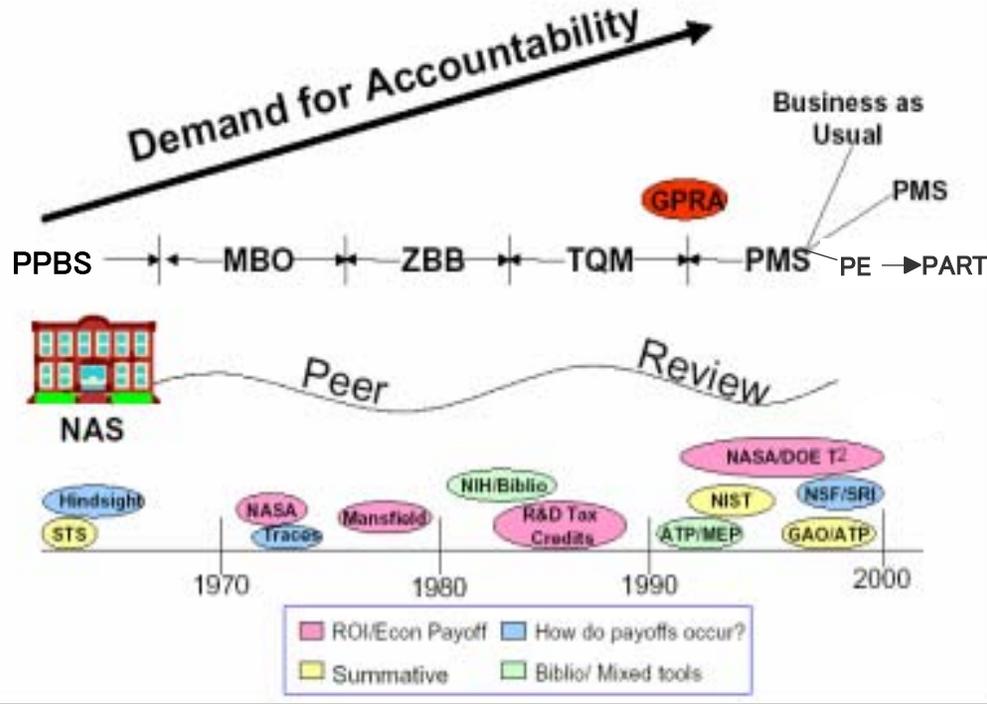
評価項目	評価基準
目的の必要性 / 妥当性 想定する状況に対する適合性 代替案に対する相対的重要性	状況に対応した目的設定になっているか 想定される代替案より重要か 想定される代替案より早期に取り組むべきか (優先性)
他の項目は「従属型プロジェクト」と同様	

主要国における評価の特徴

国名	特徴的な仕組みと評価理念
フランス	<ul style="list-style-type: none">・「ギャランター(各セクターからの代表者)」制度(評価における代表民主制)・社会の中でのディベートに一度委ねた後に意思決定
UK	<ul style="list-style-type: none">・「契約」に基づく研究活動と、資金配分に見合った価値基準(value for money)による評価・評価結果の資金配分への直接的反映
ドイツ	<ul style="list-style-type: none">・下部機関への「信託」と下部機関による自律的運営、3%ルールによる規律の保持・支援的評価システム
EU	<ul style="list-style-type: none">・プログラムの追跡評価による見なおし(フレームワーク計画)・社会経済性評価の発達
アメリカ	<ul style="list-style-type: none">・多重に設けられたチェックアンドバランスの評価の枠組み・ピアレビューとメトリックスの発達・循環型評価システムの導入(GPRAからPARTへ)
韓国	<ul style="list-style-type: none">・一元的統合評価制度と評価支援体制の強化・Authorization と Appropriationに相当する段階評価

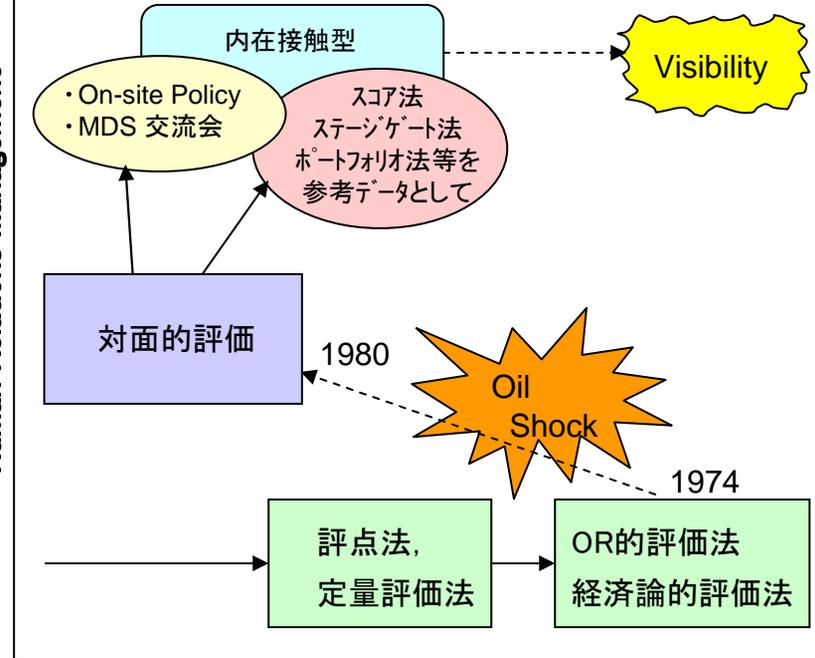
評価理念の変遷

Figure 1: Milestones in the History of Research Evaluation in the U.S.



Human Relations Management

日本の民間企業における評価



Scientific Management

David Roessner, "Outcome Measurement in the United States: State of the Art", AAAS, Boston, (2002.2)

- ・初期最適化から学習型へ
- ・計量的手法の絶対化から相対化へ
- ・定性的手法の進化と復権

Program Assessment Rating Tool(PART)

- ・行政改革のビジョン(選挙公約, 2000年)
 - －Citizen-centered, not bureaucracy-centered
 - －Results-oriented
 - －Market-based

- ・The President`s Management Agenda (OMB, 2001年8月)
行政機関のパフォーマンス向上のための評価項目と評価基準の設定

[評価項目1]

- －Human Capital
- －Competitive Sourcing
- －Financial Management
- －E-Government
- －Budget and Performance Integration

[評価項目2]

- －Program Purpose and Design (20%)
- －Strategic Planning (10%)
- －Program Management (20%)
- －Program Results (50%)

[評価項目3]

- －各5～10項目程度のサブ項目

[評価対象]

- －プログラム
- －対象を順次拡大
- －long-termとannualに分ける
- －プログラムを7種に分類
 - Competitive Grant P
 - Block/Formula Grant P
 - Regulatory Based P
 - Capital Assets and Service Acquisition P
 - Credit P
 - Direct Federal P
 - R&D P

[評価基準と評定区分]

- －目標に対する達成基準(Y/N)
- －「結果」については
達成度により4区分

[評価者]

- －evidenceを付してまず自己評価
- －OMBと協力機関でその結果を確認

[評価のまとめ]

- －状態と変化に分ける
- －プログラム毎に評価を集計
- －さらに機関毎に集計
- －青(すべてのプログラムが基準以上),
黄([青][赤]以外),
赤(一つでも顕著な欠陥があれば)

で表示

- ・Management Scorecard (Baseline Evaluation 2001年12月)
FY2003(2002年10月からの予算)のための予算要求に反映
以後毎年対象を拡大するとともに、対象プログラムによっては四半期毎にパフォーマンスの変化を測定する

「評価論」からみた現状と問題点

- ・評価制度（法律）の制定が先行し、パブリックマネジメントの実態が追いついていない
 - －評価制度の導入から始めたことの不合理
 - －評価制度の全体設計が不十分，未成熟
 - ・評価活動の全体像に対する理解が圧倒的に不足している
 - －評価の理念，評価対象とその概念構造，評価の方法論，評価人材の役割と期待される能力
 - －「評価」を支える基盤的知識
 - ・評価の多様性と取組みの難易度に関する認識の欠如
 - －従属プロジェクト，機関評価 ≪ 独立プロジェクト，プログラム，施策・政策
 - －科学技術性評価 ≪ 社会経済性評価
 - －中間，直後評価 < 追跡評価 < 事前評価
 - ・評価を担える人材や組織・機関が育っていない
 - －評価者，評価運営者，評価のための調査・分析専門家
 - －評価担当部署，評価支援外部機関
 - ・評価インフラの整備の遅れ  欠如
 - －評価データベースの欠如
- ・評価現場での混乱、しかし
 - －評価担当部署の努力，評価支援外部機関のスキルアップ
 - －より困難な課題が次々と顕在化