総務省政策評価に関する統一研修
平成31年2月13日　さいたま会場

政策評価の総論
―政策評価とEBPMの現在―

新潟大学法学部教授　南島和久

この講義のポイント

1. 評価の目的
2. 評価とは何か
3. プログラムとは何か
4. 業績測定に関する補足
5. EBPMと政策評価の関係
6. 実例

1. 評価の目的

purpose of evaluations

評価の目的(1)

評価の目的

学　習

説明責任

効果的活動

効率的活動

有効性

効率性

説明責任

評価の目的（2）

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）

第１条「この法律は、行政機関が行う政策評価に関する基本的事項等を定めることにより、政策の評価の客観的かつ厳格な実施を推進しその結果の政策への適切な反映を図るとともに、政策の評価に関する情報を公表し、**もって効果的かつ効率的な行政の推進に資するとともに、政府の有するその諸活動について国民に説明する責務が全うされるようにすることを目的とする。**」

評価の目的(3)

- 形成的評価と総括的評価(Scriven)

① **<u>形成的評価(Formative evaluation)</u>**

「コックが調理中にスープを味見すること」

・事業や活動の開発・形成

・継続的改善

・立案、実施段階で行う評価

② **<u>総括的評価(Summative evaluation)</u>**

「客が出来上がったスープを味わうこと」

・事業や活動の成果の把握

・プログラム継続・縮小・拡大の判断

・一定期間経た段階で行う評価

相互に補完し合うことで、より効果的なプログラム形成につながる

評価の目的(4)

- 評価情報を何のために使うのか
  - o 事業や活動の改善を目的としたもの(**形成的評価**)
  - o 介入のアカウンタビリティの確保を目的としたもの　(**総括的評価**)
  - o 科学的な知識・知見を生み出すことを目的としたもの

- その他にもあるが、重要なことは、評価の「目的」をまず明らかにすること。それに合わせて評価方法(誰が、いつ、どのような手法で行うのか)を選択する必要がある。

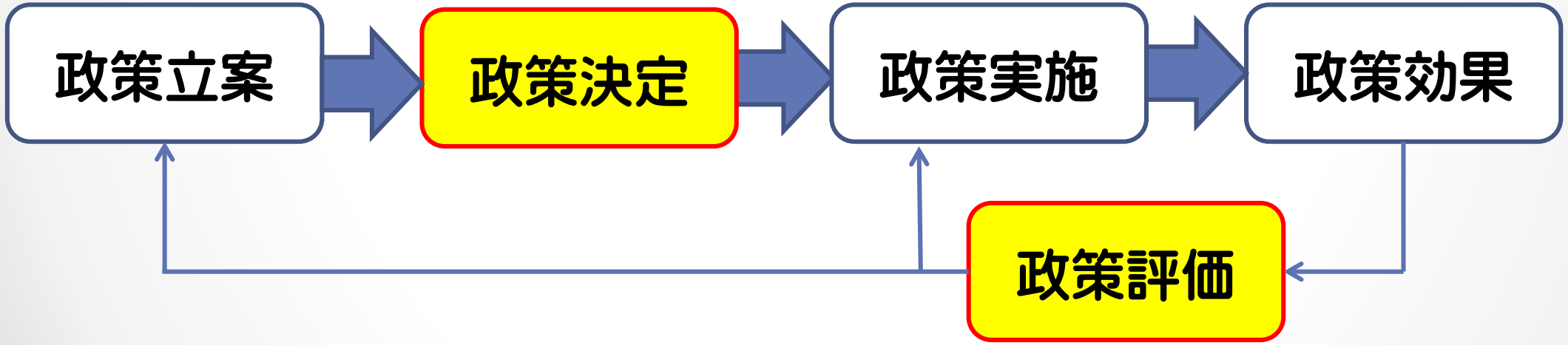
2. 評価とは何か

definition of evaluations

政策過程と政策評価

〇政策評価は、「政策立案」「政策決定」「政策実施」の前提となる（を支援する）「情報提供活動」。

→**重要**:「政策評価」は「政策決定」と異なる。



- 政策実施の改善へのフィードバック ➡形成的評価
- 政策立案へのフィードバック ➡総括的評価

評価の定義

（広義）「ものごとの利点、価値およびその意義を体系的に明らかにすること」（Scriven 1991:139）

（狭義）「社会プログラムの働きと効果に関する情報の収集、分析、解釈、伝達を目指す社会科学的活動」（「評価は、プログラムを継続すべきか、改善すべきか、拡張すべきか、それとも縮小すべきかの決定を支援するすること、新しいプログラムや先駆的事業の有用性をアセスメントすること、プログラムの運営や管理の効果性を高めること、プログラムのスポンサーからのアカウンタビリティの要求に応えること、といった様々な実用的目的のために実施される。」）」（Rossi et.al. 2004:2）

3つの評価方式

政 策 分 析 **Policy Analysis**	事業評価方式（公共事業評価、規制影響評価）、PPBS
プログラム評価 **Program Evaluation**	総合評価方式（ GAOのプログラム評価、行政評価局調査の一部）
業 績 測 定 **Performance Measurement**	実績評価方式（府省の自己評価）、GPRA、独立行政法人評価、地方公共団体の行政評価

※1　PPBS：Planning, Programing, and Budgeting System
※2　GAO:Government Accountability Office
※3　GPRA:Government Performance and Results Act,1993

政策分析（1）

政策分析

コンセプト：政策の決定は合理的であるべき　→「合理的選択モデル」

①枠組みの整理 → ②代替案の列挙 → ③効果の予測等 → ④最適解の選択

○　政策分析の典型例は、公共事業評価　＝費用便益分析（B/C＞1）

→　複数の路線整備候補のうち、どれを整備するか、しないか？

整備をする場合、どの路線を優先するか？

政策分析（2）

●事業評価方式の定義（「政策評価に関する基本方針」（平成17年12月16日閣議決定）別紙）

- 「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」

- →「公共事業」などの巨額の費用を要する政策の「事前評価」が典型

プログラム評価（1）

プログラム評価

コンセプト：「アウトカム」を多面的に検証する

→「プログラム」への注目

○「プログラム」とは何か？

- 組織活動のコンセプト
- 組織活動の機能的単位
- 複数の事務事業の複合
- 組織活動のルーティン
- ロジックモデルで表現　など

プログラム評価（2）

●総合評価方式の定義（「政策評価に関する基本方針」別紙）

- 「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」

- →政策実施を踏まえた本格的な「事後評価」（ex.行政評価局調査の一部）

※ 行政評価局調査：行政評価局調査とは、総務省行政評価局が府省とは異なる立場から、複数府省にまたがる政策や各府省の業務の現場における実施状況を実地に調査し、各府省の課題や問題点を実証的に把握・分析し、改善方策を提示するもの。同調査には、政策評価法に基づく「政策の評価」と総務省設置法に基づく「行政評価・監視」がある。

業績測定（1）

「業績測定」

コンセプト:「計画」「目標」「指標」の進捗を管理する

＝ 既存の「計画」「目標」「指標」を前提とする評価

ex.自治体の総合計画や個別計画の評価（行政評価、事務事業評価）、
独立行政法人の中期目標・中期計画を前提とする評価（独法評価）、
国の府省の事前分析表を前提とする評価（目標管理型評価）など

⇒ ただし、「計画」「目標」「指標」が合理的なものであるとは限らない。
（※ 「業績測定」「実績評価方式」「目標管理型評価」の概念には
ズレがある点に注意）

業績測定（2）

●実績評価方式の定義（「政策評価に関する基本方針」別紙）

- 「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、政策の目的と手段の対応関係を明示しつつ、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組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」

- →計画に掲げられた「指標」等により、政策の実施によって得られた成果
  ＝「達成度合い」を事後的に評価。

３. プログラムとは何か

program evaluation

プログラム（Program）とは

- 「プログラム」とは
  - プログラムとは、ある社会目的を達成するための一連の活動群を指し、それらの活動を実施に導くルール、制度、組織体制や人材を含む取り組み全体を指す。
  - 特定の組織が主体となって実施する場合もあれば、複数の組織・関係者が協力して実施する取組みもある。

プログラム評価の定義

- プログラム評価とは社会調査の方法を活用し、社会プログラムによる介入の有効性を体系的に調査し、評価を行うこと。その評価は、プログラムを取り巻く政策的・組織的な文脈を考慮した方法で行われるもので、社会状況を改善するための活動の情報源となるものである（Rossi, et. al.2004: 16）

- 評価とは、プログラムや政策の改善に寄与するための手段として、明示的または暗示的な基準と比較しながら、プログラムや政策の実施やアウトカムを体系的に査定することである（Weiss1998:4）

プログラムセオリーとは

- プログラムセオリー（Program Theory）
  - 「プログラムの実施と成果との間に存在する媒介メカニズム」（Weiss）
  - プログラムの作戦や活動内容を決定するにあたり活用できる考え方であり、「特定の成果を得るためにはこういう活動が必要であろう」という「仮説」である。
  - その「仮説」が有効か、有効にするにはどうしたらよいか、を問う。
  - 「インパクトセオリー」と「プロセスセオリー」の２つが含まれる。

【プロセスセオリー】　【インパクトセオリー】

Rossi, Lipsey & Freeman, 2004: 140

ロジックモデル

- ロジック・モデルとは何か？
  →プログラムを表現する汎用的枠組み（世界標準）。

- なぜ、ロジック・モデルが必要なのか？
  →プログラムの運営方法やプログラム・プロセス（実施過程）を明らかにする。
  →社会への介入のメカニズムを明らかにすることで体系的評価を可能とする。
  →目的の設定、指標の検討を行うときの枠組みとして活用できる。
  →ステークホルダー間で、プログラムのコンセプトを共有するために活用できる。

- ロジック・モデルの構成要素は何か？
  →4要素でプログラムを表現（「インプット」「アクティビティ」「アウトプット」「アウトカム」）

ロジックモデルのイメージ（1）

インプット（投入資源等）

- 健康診断の通知等（現状の把握）
- ダイエット等の対策（目標・指標設定）
- プログラムの準備
  - ・運動用品等の購入
  - ・サプリメント購入

→

アクティビティ（具体的活動）

- 摂取カロリーの縮小
  - ・食事量を減らす
  - ・カロリーを抑える
  - （具体的手段A）
- 消費カロリーの拡大
  - ・有酸素運動をする
  - ・基礎代謝を増やす
  - （具体的手段B）

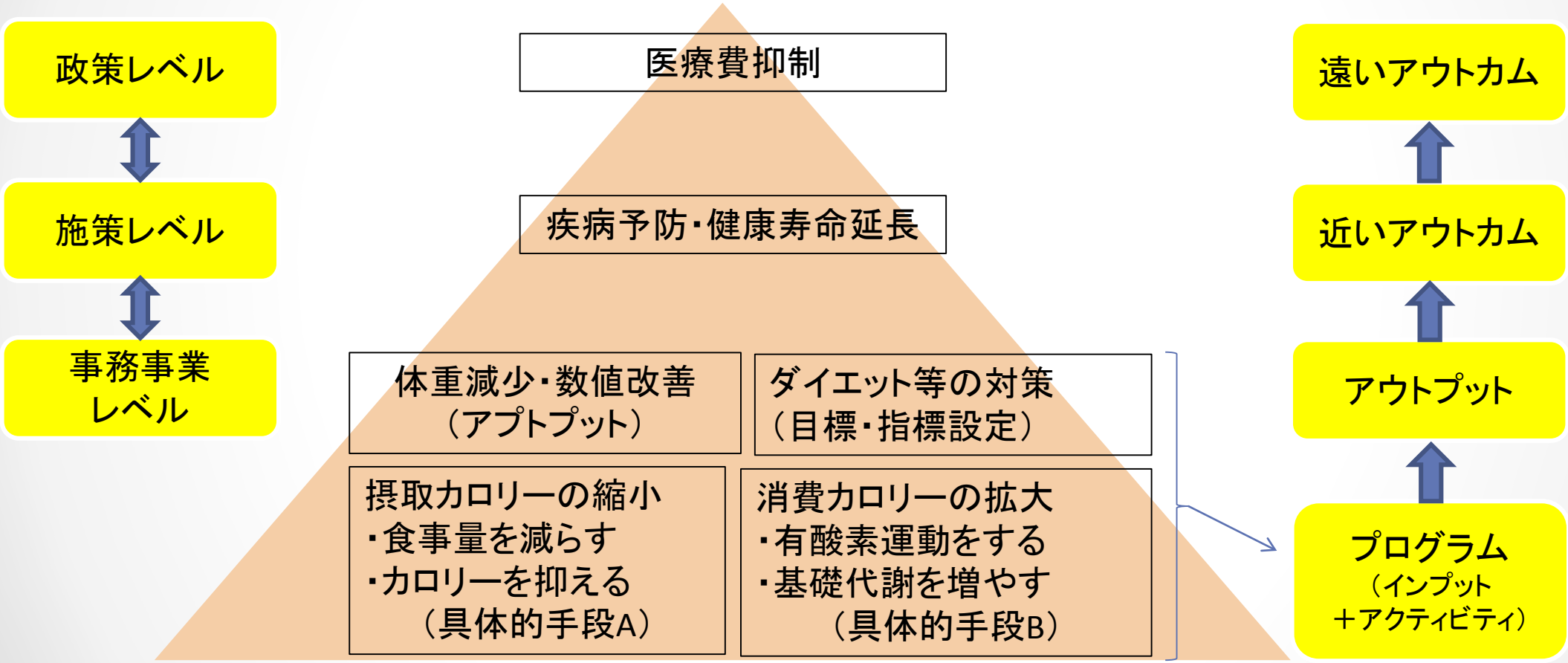
→

アウトカム等（結果・成果等）

- 体重減少・数値改善（アプトプット）
- 疾病予防・健康寿命延長（中間アウトカム）
- 医療費抑制（遠位アウトカム）

↑プログラムの定量的な管理が行われ、アウトプットやアウトカムの制御へ向かう。

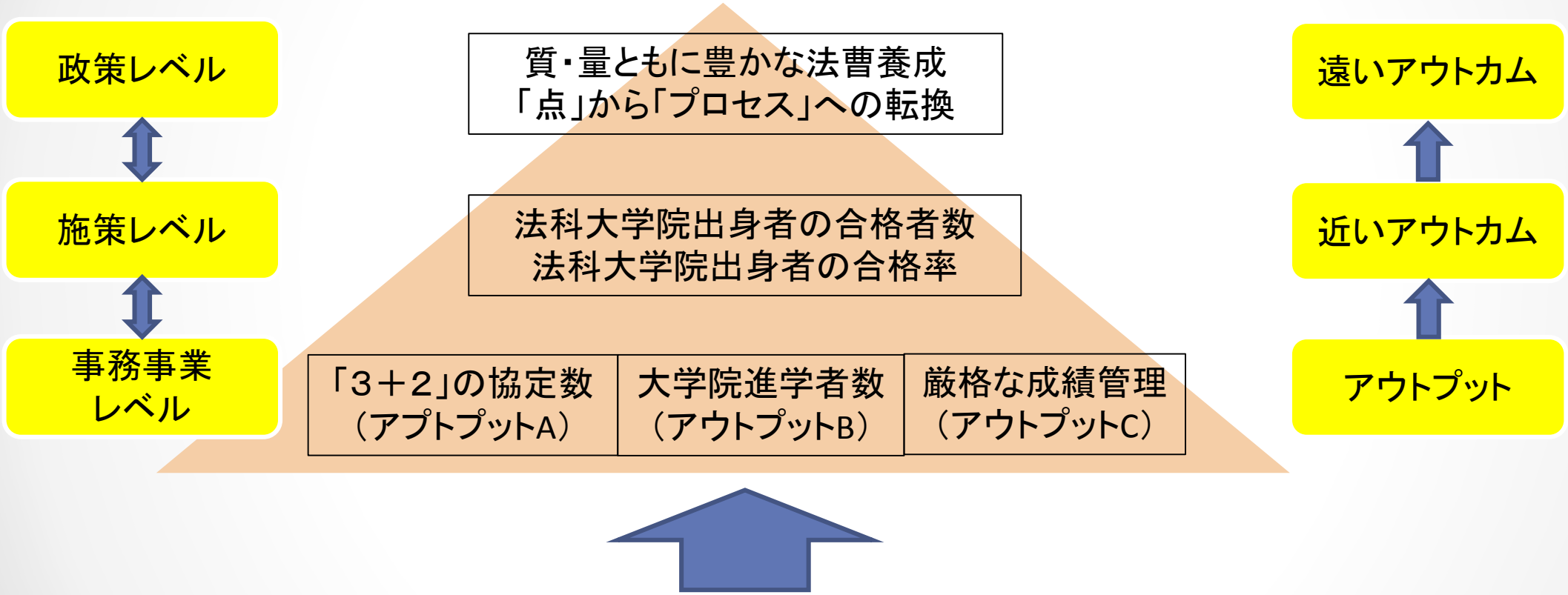
ロジックモデルのイメージ（2）



法曹養成プログラムの例（1）

インプット	アクティビティ	アウトプット	アウトカム
法科大学院への各種予算措置、学生の経済的負担の軽減策等	学部3年（基本7法）+大学院2年（答練等）→最短5年の一貫教育（「3＋2」）	最終合格数（1,525人／政府目標は1,500人以上、合格率29.11%）、うち法科出身者数（予備試験336人、法科1,189人）、法科出身者合格率（24.75%（予備試験合格率77.6%））など（平成30年）	「点」から「プロセス」へ。質・量ともに豊かな法曹養成。

法曹養成プログラムの例（2）



国や個別の大学での予算措置、中期目標や中期計画での目標設定、
学生の経済的負担軽減、教員等人員の拡充、学習環境の整備など

評価の5階層～システマティックアプローチ～
（Rossi, Lipsey and Freeman, 2004: 80）

プログラムのコストと効率性の評価
Assessment of Program Cost and Efficiency

プログラムのアウトカム／インパクト評価
Assessment of Program Outcome/Impact

プログラムのプロセスと実施の評価（プロセス評価）
Assessment of Program Process and Implementation

プログラムのデザインとセオリーの評価（プログラム理論評価／セオリー評価）
Assessment of Program Design and Theory

プログラムのニーズ評価
Assessment of Need for the Program

プログラムの三つの階層

プログラム・セオリー　制度設計

制度設計の問題
⇒社会環境が変化した。
⇒コンセプトが生煮えだった。
⇒じつは目標が複数存在する。

プログラム・プロセス　体制整備

体制整備の問題
⇒実施中にプロセスが毀損した。
⇒適切な修正ができなかった。
⇒関係するアクターが多すぎる。

プログラム・アウトカム　効果発現

効果発現の問題
⇒十分な効果が発現していない。
⇒明確に効果を把握できない。
⇒「失敗だ」ということができない。

4. 業績測定に関する補足

…

performance measurement

業績測定の定義

- サービスやプログラムの結果（アウトカム）および効率を定期的に測定すること（Hatry1999）

- インプット・作業（活動）・アウトプット・アウトカム・効率（生産性）などの一連のデータを定期的に収集し報告すること（U.S.GAO 1992）

- プログラムの実績、とくに事前に設定した目標や基準の達成度合いについての体系的で継続的なモニタリング（U.S.GAO 2012）

業績測定とプログラム評価の比較

①業績測定の対象範囲は、プログラムの「プロセス」から「アウトカム」にかけての範囲。

②業績測定の目的は、プログラムの集合やプログラムを所管する機関全体の「有効性」と「効率性」（プログラムに限定されない）。

③プログラム評価と業績測定は目的や役割が相違（「網羅・悉皆型」と「深掘り型」）。

④プログラム評価と業績測定の関係は相補関係（業績測定の結果を踏まえてプログラム評価の対象を選ぶ、プログラム評価の結果を踏まえて業績測定の指標を設定するなど。）

Nielsen&Hunter[2013], Hatry[2013]

項目	業績測定	プログラム評価
タイミング	定期的・継続的	非定期的・時宜的
評価対象（数）	多数（しばしば全数）を対象に、一律	少数を対象に、個別
情報の深さ	表層的な業績データ	因果関係の追究
実施体制	部内・自己完結型	専門・第三者型

指標に求められるもの

●妥当性のある指標を設定すること

- 業績測定において妥当性とはシンプルに「測定すべきものを測定していること」。
- ロジック・モデルに基づいて指標を設定することが推奨されている。

●目標値は明確に設定すること

- 根拠となるデータに基づいて、意味が明らかで曖昧さのない数値を。
- 目標達成の難しさが異なれば、達成度合いの評価も異なる。

●実績値と目標値は的確に比べること

- 最終的な水準とそこまでの変化のどちらに着目するか、など的確に選択。

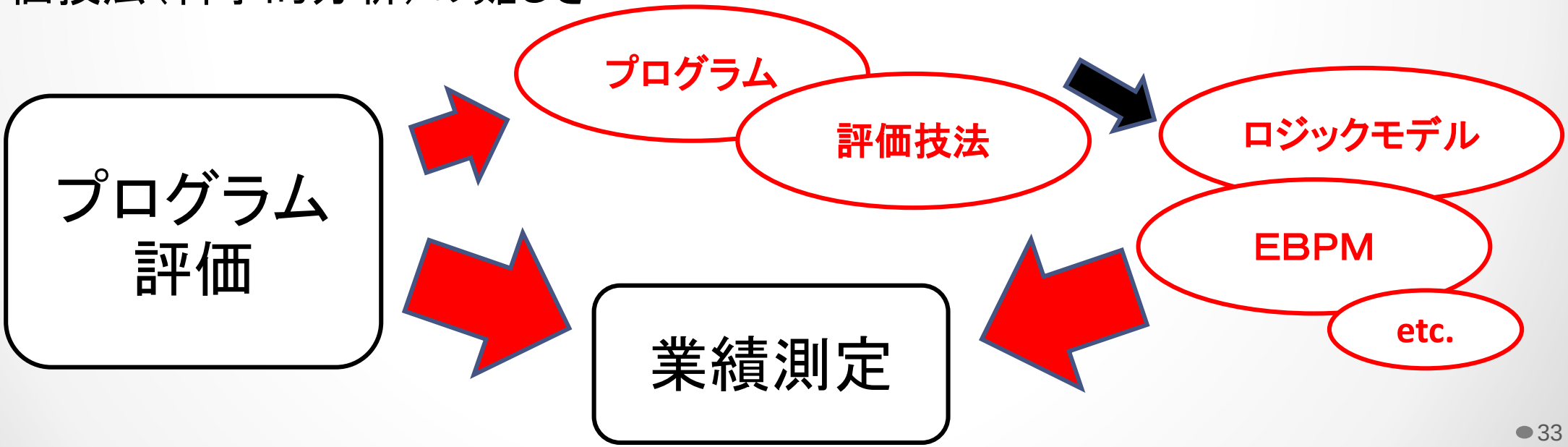
●指標間・プログラム間の比較、集計・集約は丁寧に行うこと

- 比べられないものは比べないこと（例. フロー指標とストック指標の直接比較は禁物）。

小野達也[2018]評価士養成講座テキストより

業績測定とプログラム評価の関係

- プログラムの難しさ
  ①プログラムは、しばしば所管する組織の枠組みを越えてしまう。
  ②プログラムは、しばしば単年度予算の枠組みを超えてしまう。
  ③外部要因を徹底して排除したり、プログラムを明確に定義することは難しい。
- 評価技法（科学的分析）の難しさ

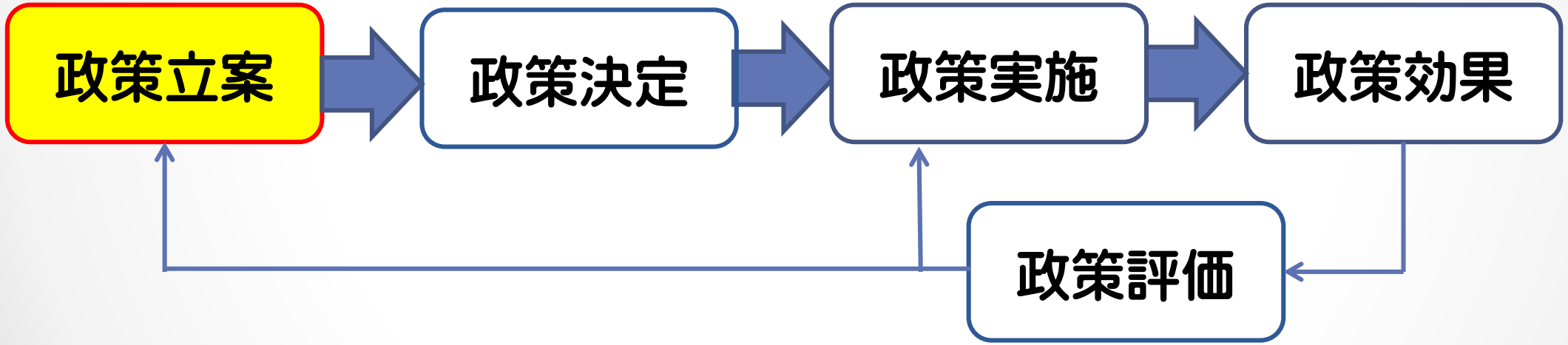


５．政策評価とEBPMとの関係

relationship between evaluation and EBPM

EBPMとは何か？

〇EBPM（evidence-based policy making）は、「政策立案」に関係する。



EBPMの3つの要請

① <u>制度趣旨明確化の要請</u>
（曖昧から明確へ）

② <u>定量的指標管理の要請</u>
（主観から客観へ）

③ <u>科学的知見活用の要請</u>
（エピソードからエビデンスへ）

→ 政策の合理性の向上

EBPMの3つの潮流

① <u>経済政策の根拠基盤の整備</u>

（統計改革推進会議⇒EBPM推進の体制整備）

② <u>医療教育分野における蓄積</u>

（学問的蓄積⇒エビデンス階層や科学的知見）

③ <u>政策評価における科学主義</u>

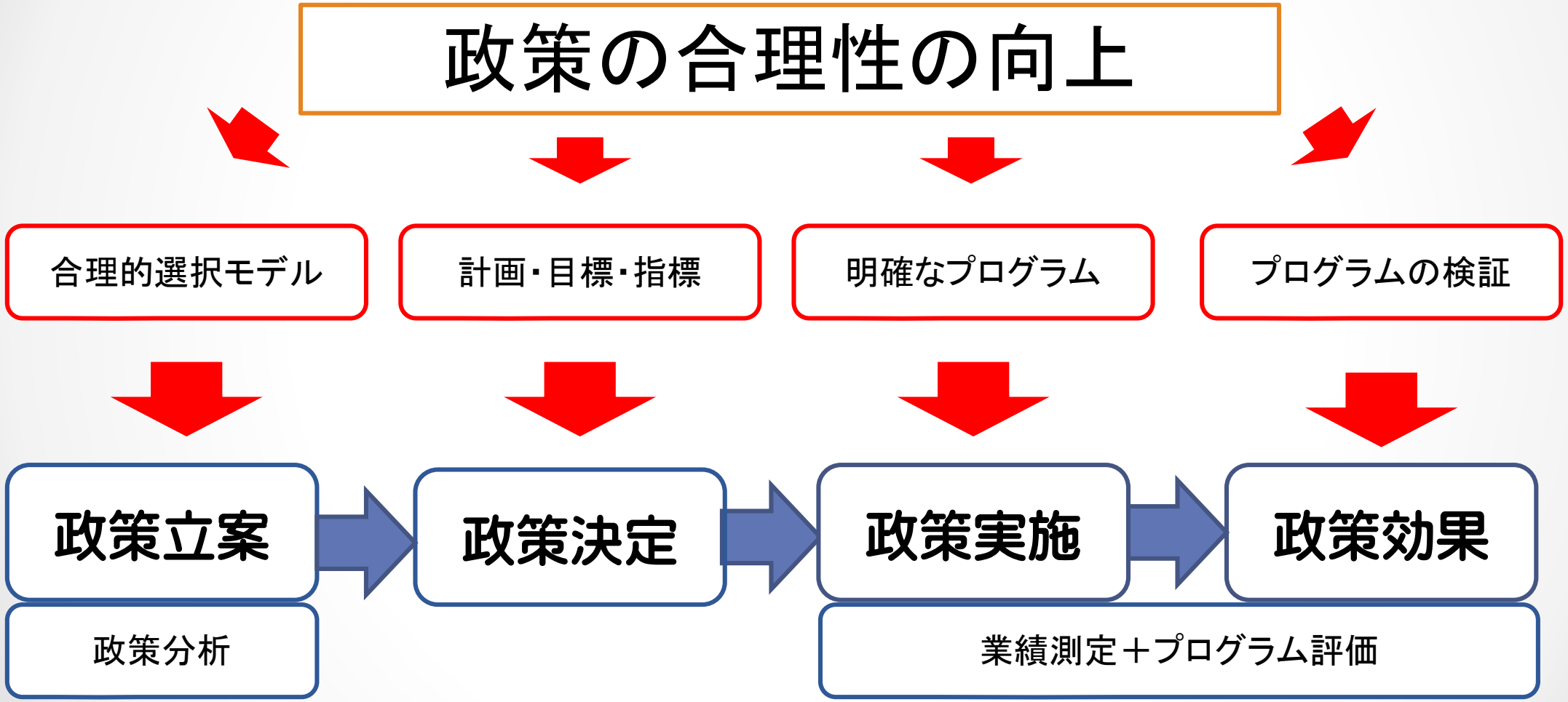
（科学主義vs.実用主義⇒政策評価の高度化）

EBPMと政策評価の関係（1）

① 「政策分析」の骨格である「**合理的選択モデル**」と親和的。
② 「業績測定」の前提となる「**計画・目標・指標**」と親和的。
③ 「プログラム評価」の「**プログラムの明確化**」と親和的。
④ 「プログラム評価」の「**アウトカム検証の技法**」と親和的。

政策評価をスパイラルアップさせる契機として期待？

EBPMと政策評価の関係（2）



業績測定とEBPMの関係

① <u>ロジックモデルの導入</u>

⇒　政策のコンセプトの明確化

② <u>効果検証技法の洗練</u>

⇒　データ・エビデンスの活用

<u>**プログラム評価への接近＝「政策評価の重点化」の推進？**</u>

6. 実例

…

practices

『愛結び』のビッグデータ活用（1）

- 内閣府の少子化対策交付金効果検証の企画分析委員会の関係で取材。取材先は愛媛県。
- 愛媛県の少子化対策事業として、県法人会連合会が受委託する事業。運用は「えひめ結婚支援センター」。
- 「愛結び」（婚活支援システム）を中核としたビッグデータの活用によるマッチング支援。
- 平成29年度は、県単独予算（約1千万円）＋内閣府地域少子化対策重点推進交付金（4千700万円）で実施。
- 15県1市に同様のシステムが広がっているところ。

『愛結び』のビッグデータ活用（2）

- 様々なデータを蓄積。専門家による「婚活支援ビッグデータ活用研究会」において分析。データに基づく議論。
- とくに研究者に列挙アルゴリズムの世界的権威が参加。データ分析に協力。同システムは総務省の「地域情報化大賞2015特別賞」を受賞。
- 参加者が失敗すればするほど、当該行動履歴の分析により、成果が上がる（マッチングに貢献する）。

『愛結び』のビッグデータ活用（3）

- 明らかな成功事例。今後、どのようにパッケージ化し、横展開につなげていくことができるのか？

  ⇒ ① 内閣府側から見た**「プログラムの汎用化の課題」**

- 愛媛県の取り組みはまだ完成していない（例えば、県の年間登録者数は800件程度で頭打ち。サンプルデータが増加すればマッチング率が向上するかもしれない。まちづくりとどのように結びつけるのか。少子化につなげていくためにはどうすべきか、など）。今後の展望としてプログラムをどのように深化させていくことができるのか？

  ⇒ ② 愛媛県側から見た**「プログラムの高度化の課題」**

IoT健康サービス事業（1）

- 総務省行政評価局のEBPMに関する実証的共同研究の関係で取材。取材先は新潟県見附市。
- 平成21年開始の9自治体によるスマートウェルネスシティ首長研究会が事業の母体（新潟県見附市、福島県伊達市、新潟県新潟市、三条市、岐阜県岐阜市、大阪府高石市、兵庫県豊岡市、千葉県浦安市、栃木県大田原市、岡山県岡山市）。
- 企業・大学などによるコンソーシアムを組成。
- 内閣府（地方創生）の「健幸長寿社会を創造するスマートウェルネスシティ総合特区」（平成29年指定）の事業。

IoT健康サービス事業（2）

- 目標は、「健康づくりの無関心層を含む住民の行動変容」の促進、「持続可能な先進予防型社会」の創造、「地域活力の沈下」の防止、「地域活性化」への貢献等。
- 具体的には活動量計・体組成計でデータ収集。無関心層への訴求とヘルスリテラシーの向上。科学的・客観的で相互比較可能なエビデンスに基づく評価を志向。
- 収集したデータに基づき行動分析を行い、医療費抑制効果・経済波及効果検証やビジネスモデルを検討。

IoT健康サービス事業（3）

- 本事業は、総務省（テレコム）のIoTサービス創出支援事業の一環。行政評価局のEBPM共同研究では、EBPMの推進に向き合いつつ、ロジックモデルの組み立てと、タテ展開、ヨコ展開に関心。詳しくは報告書（公表予定）を参照。
- ①「プログラム汎用化の課題」と②「プログラム高度化の課題」は該当。後者は例えば、市の予算の制約による募集人数の制約の課題、経済波及効果・地域経済還元効果の把握の課題、医療費抑制効果の見える化の課題、参加者の固定化の打破（無関心層への訴求）の課題などがある。

むすびにかえて

…

conclusion

まとめ

- EBPMは政策評価を高度化するためのオプション。
- EBPMは、最広義では、「合理性を高めるもの」と理解できる。その意味では政策評価と方向性は共通。業績測定の高度化にも活用可能（なお、ICT関係とEBPMは親和的）。
- 鍵は、①「プログラムの汎用化」と②「プログラムの高度化」。
- 「所管の壁」「予算の壁」により、「プログラム」の切り取り方は難しい。

⇒　組織内外のコミュニケーションの拡大と充実を図りたい。

参考文献

秋吉貴雄・伊藤修一郎・北山俊哉（2015）『公共政策学の基礎』有斐閣。
足立幸男（2009）『公共政策学とは何か』ミネルヴァ書房。
石橋章一朗・佐野亘・土山希美枝・南島和久（2018）『公共政策学』ミネルヴァ書房。
今村都南雄（1997）『行政学の基礎理論』三嶺書房。
――――――・武藤博己・佐藤克廣・沼田良・南島和久（2015）『ホーンブック基礎行政学　第３版』北樹出版。
大橋洋一編著（2010）『政策実施』ミネルヴァ書房。
佐藤竺監修、今川晃・馬場健編著（2009）『市民のための地方自治入門　新訂版』実務教育出版（とくに拙著「第12章　自治の課題『評価』」）。
南島和久（2007）「府省における政策評価の中立性および客観性：グリッド／グループ文化理論に基づく考察」（『法学志林』104（4））。
――――（2011）「府省における政策評価と行政事業レビュー：政策管理・評価基準・評価階層」（『会計検査研究』（43））。
――――（2013）「NPM・行財政改革と大学評価」（広田照幸他編『組織としての大学（シリーズ大学第６巻）』岩波書店）。
――――（2015）「政策評価の概念とそのアポリア：分析・評価・測定をめぐる混乱」（『評価クォータリー』（33））
――――（2016）「米国のGPRAMAにみる制度改革への視座：日本への示唆と業績マネジメント」（『評価クォータリー』（38））
――――（2017）「行政におけるエビデンスとアウトカム：自殺対策の評価からの考察」（『季刊行政管理研究』158）。
――――（2017）「行政管理と政策評価の交錯：プログラムの観念とその意義」（『公共政策研究』（17））。
――――（2018）「自治体における評価制度の見直しの視点：成果指標を中心に」（『評価クォータリー』（45））。
西尾勝（1990）『行政学の基礎概念』東京大学出版会。
益田直子（2010）『アメリカ行政活動検査院：統治機構における評価機能の誕生』木鐸社。
松田憲忠・岡田浩編著（2018）『よくわかる政治過程論』ミネルヴァ書房（とくに拙著「第３部　X　政策評価」）。
竜慶昭・佐々木亮（2004）『「政策評価」の理論と技法（増補改訂版）』多賀出版。
山谷清志（2012）『政策評価』ミネルヴァ書房。
山谷清志編（2010）『公共部門の評価と管理』（晃洋書房、2010年）
Hatry,P.（2006）Performsance Measurement; Getting Results:2nd.ed.,Urban Institute.（上野宏・上野真紀子訳『政策評価入門』東洋経済新報社、2004年）
Newcommer,K.E and Harry P. Hatry, Joseph S.Wholey（2010;2015）Handbook of Practical Program Evaluation, Jossey-Bass.
Nielsen,S. and Hunter,D. （2013）　Performance Management and Evaluation: New Directions for Evaluation, Jossey-Bass.
Scriven,M.（1991）　Evaluation Thesaurus,4th ed. ,Sage.
Rossi,P.,Lipsey,M. and Freeman,H. （2004）Evaluation 7th ed.,Sage.（大島巌他訳『プログラム評価の理論と方法』日本評論社、2005年）
WeissC.H.（1997） Evaluation: Methods for Studying Programs and Practices:2nd.ed,printice Hall.（佐々木亮監修『入門　評価学』日本評論社、2014年）