

政策評価制度は何を解決するのか 制度運用の実態と理論の考察を通じて（要約）

東京大学大学院修士課程 政治専攻
公共政策専修コース ・学籍番号 26164
益田 直子

本稿は紙幅の都合上、論文の目次、本文要約、本文一部の抜粋から構成されている。抜粋箇所は、評価研究について知りたい方を想定して第2章の評価研究における議論の変遷を、そして、評価の実践の現場における今後の議論の高まりを予想して第6章第3節の評価の利用に関する議論（一部）を掲載している。ご了承ください。

<目次>

- 序章
- 第1章 行政不信・財政問題の要因
 - 1.1. 行政不信の要因
 - 1.2. 財政問題と評価
 - 1.3. 第1章まとめ
- 第2章 評価研究における議論の変遷
 - 2.1. 政策評価の起こり
 - 2.2. 政策評価における議論の展開
 - 2.3. 最近の変化
- 第3章 日本において行政システムの合理化制度が現れる背景
 - 3.1. PPBS (Planning-Programming-Budgeting System (計画策定・プログラム作成・予算編成システム))
 - 3.2. 政策評価
 - 3.3. PPBS と政策評価の比較
- 第4章 問題提起
- 第5章 実証分析
 - 5.1. 概要
 - 5.2. 政策評価の作業過程・利用状況及び意思決定過程への影響
 - 5.3. 政策評価の利用を促すインセンティブ
 - 5.4. 政策評価によって産出される情報とその利用
 - 5.5. 第5章まとめ
- 第6章 評価研究・先行事例での取り組み
 - 6.1. 政治（意思決定過程）と政策評価の関係
 - 6.2. 政策評価利用のインセンティブ創出をめぐる議論
 - 6.3. 政策評価の結果の利用
 - 6.4. 第6章まとめ

結論

<要約>

本論文で取り上げる政策評価制度は、従来の行政が法律の制定や予算の獲得に重点を置き、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を軽視しがちであった状況を問題と捉え、政策の効果について評価を行い行政機関の企画立案作業に反映させる仕組みを強化することを目的に、2001年1月に導入された。具体的には、「国民に対する行政の説明責任の徹底」、「国民本位の効率的で質の高い行政の実現」、および「国民的視点に立った成果重視の行政への転換」の3つを目的としている。制度導入後4年目を迎えるようである現在の、日本の評価研究においては、諸外国の先進的評価手法や理論の説明、または個別行政機関の取り組みの紹介

が中心となっており、日本の政策評価制度導入後の実態を行政の現場から調べるとともに、その目的、制度設計、又は評価研究の議論に照らして、制度の現状がどのような構造と課題を示しているのかに関する研究がほとんど存在していない。そのため、本論文では、政策評価制度は何を解決するために生まれたのか、果たしてその問題を解決できる制度なのか、制度が依拠している評価研究は解決策を用意しているのか、そして、制度設計とその運用のあり方は一致しているのか、という問いに取組んでいる。

論文の構成は、8つの章に分かれている。序章で上述の問題意識を明らかにした後、政策評価制度導入の背後にある行政不信・財政問題の要因を明らかにする（第1章）。その要因の対応策として期待される評価研究がどのような性質を持つものかについて、評価研究の起こりから現在に至るまでの議論の変遷を紹介する（第2章）そして、政策評価制度のような行政システムの合理化を目的とした制度は過去にも議論されていることから、その制度がたびたび現れる背景について検証する（第3章）。このように、第1章から第3章で論じた、政策評価制度の導入の背後にある要因の分析、評価研究の性質、そして日本で新しい行政手法が必要とされる背景から浮かび上がる課題と論点を、第4章で提示する。

制度をめぐる中心的課題は「社会環境の変化に迅速に反応することを目的とする『行政の機動性の確保』および累積赤字への対策を目的とした『歳出合理化』にいかにか政策評価制度は対応できるか」にあり、この課題に答えるために、次の3つの論点を本論文で検討すべき具体的な問いとした。第一は、閉鎖的な意思決定過程への影響である。意思決定過程における「政治家・関係団体（利益集団）・行政機関」の関係、または予算編成過程における「国会の常任委員会・与党政策部会・関係団体（利益集団）・行政機関」の関係といった既存の閉鎖的な意思決定過程に対する政策評価の関与について問うものである。第二は、評価利用のインセンティブとその影響である。政策評価による意思決定への関与および利用に関する現状を調べ、その利用を促すインセンティブはどのようなものであり、それらが評価の利用に今後与える影響について問うものである。そして、第三は、行政組織内に入ってきた行政活動上の判断に資する情報が消滅する理由とその回避についてである。既存の関係主体からのさまざまな介入を受けながらも、優先すべき受益主体の変化やリスク等に関する情報を組織の判断に資する情報として感知し、政策の立案から実施まで繋げるには何をすべきか、という問いである。

これらの論点について、まず、評価の実施府省から農林水産省、国土交通省、厚生労働省の3省を選び、また、これら実施府省の評価のメタ評価を行う評価専任組織である総務省行政評価局を取り上げ、評価制度の設計および運用の実態調査を行った（第5章）。続けて、評価研究・先行事例における議論（第6章）を通じて、制度をめぐる構造を示した上で、政策評価制度の今後の方向性について論じている（結論）。

その結果明らかとなったのは、次のことである。実施府省による評価の現状は、評価利用のインセンティブの獲得が課題となっており、その対応策として予算編成および事業採択の判断に評価を利用する取組みが進められていることから、事業評価（事前）と実績評価の利用が中心となっている。内閣府経済財政諮問会議も予算と評価のリンクの試みを始めており、「歳出合理化」への取組みは進められている。また、評価結果の公表により、政策評価制度の3つの目的のうち「国民に対する行政の説明責任の徹底」への取組みは進展を見せている。一方、制度化されているにも関わらず、総合評価、事業評価（事後）の実施や実施府省外部の主体の参画は少なく、チェック機能が十分に生かされているとは言えないことから、「行政の機動性の

¹ 本論文に関心ある方は、NaokoM012@aol.comにご連絡ください。

確保」への取組みは不十分である。また、政策評価制度の目的のうち「国民本位の効率的で質の高い行政の実現」及び「国民的視点に立った成果重視の行政への転換」への取組みも十分ではない。これらのことから、今後特に必要な取組みは、チェック機能をいかに発揮させるかにあることが分かる。

上述の制度をめぐる構造は、評価情報の産出・利用・提供・チェックのすべての段階において、行政機関の裁量が働くように制度が設計されていることと強く関係している。この裁量は、行政活動の実状に評価活動を適合させてゆく上では重要な仕掛けである。しかし、評価制度の要ともいえるチェック機能が行政機関による裁量制によって発揮されないならば、総務省行政評価局・評価委員会、および行政機関外部の主体の活動量に依存せざるを得ない。つまり、このチェック機能を各行政機関内において発揮するためには、制度上明記されていないが次のシナリオの実現が求められる。総務省行政評価局・評価委員会および行政機関外部の主体(国会、専門家、マスコミ、市民団体、そして関心を持つ個人等)など政策実施主体から切り離された主体がチェック機能を発揮し、行政機関内の評価室と相互関係を結ぶことによって、評価室は政策の実施を担う原局原課に対するチェック機能を高め、ひいては原局原課の職員による日常業務に見直しの作用が組込まれる。そして、この作用の発揮によって行政機関が社会の変化を広範な行政活動に反映する可能性を最大化する、というシナリオである。しかし、総務省行政評価局はその機能に関する各府省の合意の取りつけが不十分であり、行政機関外部の主体は評価に関心を示していない現状に鑑みて、上のシナリオが容易に進行しない様子がその始まりから示されている。

こうした制度の現状と課題を踏まえ、日本における評価研究の方向性として、日本の政治・行政研究との連動、および組織論との関わり等が求められていくと考える。「政策評価が評価であるためには、評価はそもそも行政活動に疑いの眼差しを組込む性質のものであることが認識される必要がある。世間一般の通念に疑いの目を持ち目標達成の測定に細心の注意を払い、それらのために必要な情報を求めつづけ、そして知識を基盤とした改革に寄与することが、評価の役割である」(Chelimsky, Eleanor, and William R. Shadish. 1997. *Evaluation for the 21st century*. California: Sage Publications, p. 25) という評価機能の実現への取組みが、評価研究と実務の双方に、今求められている。

<第2章 評価研究における議論の変遷(抜粋)>

本章では、政策評価に関する研究の起こりと、評価研究におけるこれまでの議論について論ずる。つまり、本論文全体の問題意識である行政不信・財政問題の要因への対応策として政策評価は、その研究においてどのような性質を持つものと論じられているのかについて説明する。

2.1. 政策評価の起こり

1965年、アメリカにおいてPPBS(Planning Programming Budgeting System(計画策定・プログラム作成・予算編成システム))および「偉大な社会」政策の導入は、政策評価の誕生の契機となった。この頃、反貧困プログラムの増大に伴い社会科学的手法は、政府活動の成功を助け、機能しないプログラムを排除するものと期待された。こうした状況から、世界の理解を図る伝統的な社会科学から、公的セクターの活動とその広範な社会へのインパクトを測る政策評価が生み出されることとなった²。

² (Haveman 1987)を参考とし、筆者がまとめている。

その後1970年代まで、評価研究はブームの時期を迎える。それまでの評価研究は社会科学者の関心に応えるものであったが、評価研究の利用者の視点が評価研究に大きな影響力を持つようになり、政治的決定や評価の設計・実行などを生み出す、公共管理活動と社会政策の統合部分となっていく(Rossi, Freeman and Lipsey 1999, pp. 11-13)。

その後、行政組織による作用の見直しの必要性の高まりに伴い、政策評価は国際的な広がりを見せている³。

この広がりの中に、日本もいる。政策評価の起こりの背景に似た状況に日本もある。公共管理の見直し(行政活動の見直し)そして財政逼迫下における政府プログラムのスクリーニングが求められている。財政赤字、薬害エイズ、旧大蔵省の過剰接待問題、外務省の機密費流用、政治献金問題、BSE発生などの事件を、新聞紙上で目にしてきた。こうした官僚の不祥事、公共政策の失敗により、行政活動への不信感は高まっている。

上述の評価研究の経緯のように、公共管理の見直しと評価のつながりは、組織運営の改善のための評価の利用という視点を生み出す。また、公的セクターの活動による社会へのインパクトを測る評価の誕生は、行政組織の外部環境との関係性の模索の始まりとも言える。

こうした動きは、日本の政策評価制度においても見られる。政策評価制度の目的は前述のとおり、国民に対する行政の説明責任の徹底、国民本位の効率的で質の高い行政の実現、そして国民的視点に立った成果重視の行政への転換であり、行政機関の機能の改善および国民の視点をより強調した行政運営が進められようとしている。具体的には、政策評価制度は、行政機関がその所掌に係る政策についてその効果を把握し、必要性、効率性、又は有効性の観点から自ら評価を行い、その評価結果を政策に適切に反映させなければならないと定めている(政策評価法3条)。そして、行政機関による評価結果の政策への反映状況の公表(政策評価法11条)および総務大臣による評価情報の所在に関する情報の提供についての措置が講ぜられることを義務づけている(政策評価法22条)。

2.2. 政策評価における議論の展開

政策評価の手法やモデルは多様化し、それをめぐる議論が展開してきた。例えば、1960年代にDonald T. Campbellは、真実を提供すること、および社会プログラムが効果を与える因果関係を明らかにすることを優先的に重視した。実験手法による分析の結果が、政策立案者やプログラム管理者によってプログラムの維持・拡張・変更の判断に利用されることを期待していた。これに対し、1970年代に入ると、それまでの理論がプログラムの改善に有用な知識を提供してこなかったことを問題とするようになった。つまり実際の政策決定過程に評価結果があまり利用されていないことを問題とし、忠実な真実の探求ではなく、利用可能な程度に正確な評価を行えば良いという考え方が生まれるようになる。そして、評価者が事前に重要な事項を特定できるということを期待せず、また、プログラムが影響を与える因果関係を最優先事項としないという考え方があらわれる。むしろ、プログラムがどのように実施されたかを説明する

³ (Love 2001, p. 442)によると次のとおり、評価関連協会の数が増加している。

1970年代 - 1980年代 3
- 1998年 10未満
- 2001年 40以上

また、International Organization for Cooperation in Evaluation (IOCE)が、AEA International&Cross Cultural Topical Interest group (I&CC TIG)によって提案されている。

ことに重きが置かれるようになる。そのため、評価手法についても実験手法のみならず、方法論の多元化が主張されるようになった。量的アプローチと質的アプローチの議論が出てきたのもこの頃である。1980年代に入り Lee J. Cronbach や Peter H. Rossi は、1960年代以降に現れたさまざまな処方箋を統合させようとした。また、一つの理論をすべての場面で利用しようとするのではなく、異なる状況と目的の下でそれに適した評価実践が意味を持つことを明らかにしようとした。そのため両者は、適切性・妥当性・許容力の観点から、さまざまな手法の長所と短所について議論をしている。この長所と短所に基づき、必要とされている情報に合う手法の選択を試みている⁴。

さらに、1993年アメリカにおける GPRA (Government Performance and Results Act) の導入をきっかけに、パフォーマンス・メジャーメントが成長するに至る。これは、目標を明らかにし、成果を測定する指標と数値目標を定め、定期的にその指標値を収集することにより、当初の数値目標がどれだけ達成されたかを評価し、現場での実施改善と意思決定に利用する仕組みである⁵。

ほか、事実と価値、客観的評価とアドボカシ - 評価など、専門家を二分する議論も生まれている (Smith 2001, pp. 292-295)。

2.3. 最近の変化

評価研究に起きた最近の変化として、次のものが挙げられる⁶。

第一に、広範な利害関係者による評価活動への参加の実施、あるいは考慮がされている。具体的には、評価対象の政策によって影響を受ける主体の参画(影響測定の精緻化とも捉えうる) (Collaborative and Participatory Approaches) 評価結果の一般化からこぼれ落ちる危険性の高い主体への配慮 (Constructivist Approach) が挙げられる。

第二に、評価の利用の向上 (Qualitative Studies, Meta Analysis, Program Theory, Advocacy (利害関係者による評価結果の受入れの重要性を強調)) が挙げられる。

第三に、組織の意思決定過程への関与を強め、その改善や変革を目的とした利用が期待されている。中には、プログラム開発者としての評価者の議論の始まりが見られる。

第四に、組み合わせられた手法の利用(学際的な視点、多様な手法を視野に入れた思考)が進められている。

第五に、評価者の専門性への懐疑が現れ、評価者の資格に関する議論が起きている (2000年 American Journal of Evaluation における議論)。

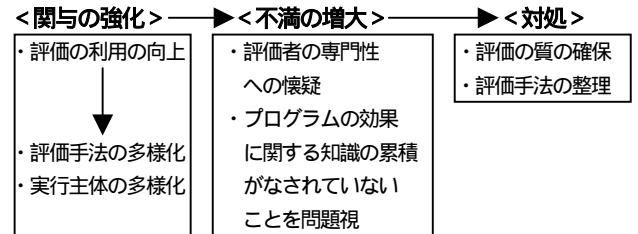
第六に、評価手法に関する知識の不足よりも、プログラムの効果に関する知識が累積されていないことを問題視する動きが現れている。

最後に、評価の質や手法への懸念が指摘されている。具体的には、評価の質の確保 (Metaevaluation) や評価手法の整理の必要性が指摘されている(例:Stufflebeam 2001, pp. 7-98)。

これらは、次のことを物語っている。つまり、評価は、政策に対する広範な利害関係者が評価活動に参画することを促すとともに、組織の意思決定への関与を強めようとしている。しかし

その一方で、評価者の専門性に対する懐疑が現れ、また、プログラムの効果に関する知識が累積されていないことへの問題意識が高まっている。これは、評価の利用の向上によってもたらされる、手法の多様化および評価の実行主体の多様化による結果と考えられる。これらの問題への対処として、評価の質の確保、および評価手法の整理が強調され始めたと言えるだろう。この関係は図2-1のように示すことができる。

【図2-1】



さらに、次のような危惧さえ感じられる。

これまでの説明にあるとおり、政策の影響への不信、その立案・執行者や執行機関への不信が契機となって、政策評価の重要性が増大している。この不信感が高まれば高まるほど、評価に期待される役割は拡大していく。そして、その拡大する期待に応えるように、評価の利用の向上や評価手法・実行主体の多様化が促進される。また、評価への期待の一方で意思決定にあまり利用されていないという認識があるため、評価が利用されるために、さらにその評価手法・実行主体の多様化は促される。

しかし一方で、評価の専門性への懐疑、および、評価によって検証されるプログラムの効果に関する発見、つまり、何が機能して、何が機能しなかったのか、そしてどのような状況でそれらが起きたのかについての知識の累積がほとんどないことへの不満が生じている。

つまり、行政活動に対する不信への対策として、評価は利用され始めたにもかかわらず、その対策としての期待感が高まれば高まるほど、評価への不満が増大するという危険を抱えている。評価への期待は、行政への期待とこの点において似ている。評価活動が行政活動の宿命と重なって見える。

評価活動も行政活動も、政策が対象とする世界のすべてを捉え、担うことはできない。双方とも、その能力を超えた範囲に踏み込み、不信や不満を生み出している。

<第6章第3節 政策評価の結果の利用(抜粋)>

ここでは第3の論点である、政策評価が提供する情報の性質およびその情報の消滅理由との関係について、評価研究の議論を通じて考える。まず、評価情報の性質に関する評価研究の議論を紹介する。評価情報の性質は、実際には評価情報の利用方法として現れてくるため、評価情報の利用の観点から分析する。次に、5.4.1.(1)で説明した日本の政策評価制度において産出が期待されている情報の性質と評価研究での議論を比較した上で、評価情報の消滅を回避する策について考察する。

6.3.1. 評価情報の利用

(1) 分類

評価研究において、評価情報の利用に関する考え方の分類として次のものが考えられてきた。まず、評価結果に基づいてプログラムを変化させたり、新しいプログラムを採択するなど、直接的で目に見える効果をもたらすための利用である「道具的利用 (instrumental use)」がなされた (Shadish, Cook and Leviton 1991, p. 53; Rossi, Freeman and Lipsey 1999, p. 432; Vedung 2000, p. 269; Owen and Rogers 1999, p. 108)。

⁴ 1960年代 - 1980年代以降の評価手法をめぐる議論については、(Shadish, Cook and Leviton 1991)を参考としている。

⁵ (Hatry 1999, p. 3)においては、パフォーマンス・メジャーメント(業績測定)の定義は、「サービス又はプログラムの結果(アウトカム)と効率性を、日常的に測定すること」とされている。

⁶ (Smith 2001, pp. 281-300)の議論を参考に、筆者の分析が加えられている。

次に、評価によりプログラムの前提・問題・メリット・限界について理解をすること、そして、主要な政策についての決定に評価の積み重ねが影響を与えることに評価情報の利用の焦点が当てられ、直接的に評価結果が行動に反映されることを目的としない利用方法である、「概念的利用 (conceptual use)」、「啓発的利用 (enlightenment use)」、または「啓蒙 (demystification)」と呼ばれる方法が主張されるようになった(以下、同義で利用されているこれらの用語は、「概念的利用」または「啓発的利用」のどちらかを論者に依拠して利用) (Shadish, Cook and Leviton 1991, p. 53; Rossi, Freeman and Lipsey 1999, p. 432; Owen and Rogers 1999, p. 108)。その他には、現状の防御または攻撃のために利用する「説得材料としての利用 (persuasive utilization)」（Rossi, Freeman and Lipsey 1999, p. 432）対話が社会問題を知的に洞察することに貢献すると考える「多様なアクター間における相互作用の利用 (interactive utilization)」（Vedung 2000, p. 274）そしてプログラムの実施・削減などの既に行われている決定を正当化するための評価である「正当化アプローチ (legitimizing approach)」（Vedung 2000, p. 275; Owen and Rogers 1999, p. 110）が考えられてきた。このように、評価情報の利用についてはさまざまな考え方が示されてきた。特に、「道具的利用」と「概念的利用」のどちらを利用するのかという議論は、現在もなお続いている。

(2) 「道具的利用」と「概念的利用」をめぐる議論

「道具的利用」と「概念的利用」をめぐるこれまでの議論について、Shadish, William R., Jr., Tomas D. Cook, and Laura C. Leviton. 1991. *Foundation of program evaluation*. California: Sage Publications. に基づき説明する。

1960年代のMichael ScrivenやDonald T. Campbellの評価利用に対する考え方は、「単純な道具主義 (naive instrumentalism)」と呼ばれる。彼らは、政策決定者とプログラム管理者が効果的なプログラムを維持または拡張するために、そして非効果的なプログラムを改善するために、評価結果が利用されると考えていた。また、そのような利用を促すような評価者の役割はほとんどないと考えていたことから、彼らは利用について広く議論することはなかった。

これに対し、1970年代にCarol H. Weissは、プログラムの失敗に関する証拠があったとしてもプログラムはそのまま残っており、逆に肯定的証拠があったとしてもそのプログラムの終了を回避できない事実を指摘した。非効果を明らかにする情報があったとしても、プログラムが一般に受け入れられている価値に合う場合やプログラムが有権者を満足させている場合などにそれは許容されてしまう状況を見てWeissは、評価研究ができることは、どのような政治的トレード・オフが存在しているのかを明らかにすることであると考えた。このことからWeissは、「概念的利用」を強調するようになる。評価の成功がその利用状況に依存している点については、1970年代よりも前にすでにSuchmanが指摘していたが、利用の詳細については議論がされていなかった。1970年代の初めになり利用に関する議論に変化が生じ、その代表者がCarol H. Weiss, Joseph S. Wholey, Robert E. Stakeであった。評価の利用の改善について、次の2つのアプローチが提案された。第一は、より熟考された「道具的利用」へのアプローチの提案である。それは、「道具的利用」は頻繁に起こることではないため、情報を利用する利害関係者のニーズを判断することによって、評価者は評価の利用を促進しなければならなかったことを表している。具体的には、Weissは政策決定者と中間管理職が抱える滞った決定に焦点を当て、Wholeyは結果の管理を望む中間レベルの官僚における「道具的利用」の促進に焦点を当てた。そして、

Stakeは利害関係者のところへ直接出向き、それらの人々の関心に耳を傾け関連情報を集めた。第二は、「啓発的利用」の提案である。StakeもWeissも両方とも、啓発の役割の増大を提言し、「道具的利用」を優先することに疑問を呈した。Weissは、「道具的利用」は漸進主義的な変化しか導かず、社会問題の起源と因果関係などについての情報こそが、時間を要するけれども将来の政策をより良くすると考えた。一方、Wholeyは、Weissほど「道具的利用」に対して悲観的ではなく、短期間で既存のプログラムが改善されることを好んだ。「啓発的利用」は、社会科学理論の発展との共存、アカデミックの社会的役割との共存を図ることができるという点において、Weissのようなアカデミックからの研究者による利用が好まれたと言うこともできる。しかし実践の場では、社会的に重要な問題への解答や、時に啓発を必要とされない制約を受けた状態での評価の実施が要求される。この点についてWholeyは、評価者は漸進主義的な利用のアプローチを適用しなければならないことを明らかにしており、彼による評価は「道具的利用」として認識されている。一方、Stakeは、直感的に理解できる経験を引き出す文章を読んだときに評価結果は利用されると考え、ケース・スタディがこれらの経験を引き出す最良の方法であると主張している。また、評価の利用者についてはそれまで、Campbellは政策決定者と中間管理職を、Scrivenは消費者を想定していたことをWeiss, Wholey, Stakeは不十分であると考え、各評価に対して評価者が利用者を特定し、その利用者のそばで仕事をすることによって有用な情報を生み出す可能性が高まると考えた。具体的には、WholeyとWeissが連邦と州の管理者から始める一方で、Stakeは、クライアント、サービス提供者、そして地方自治体職員へと、利用者の範囲を拡大させた。

次に、これらの理論家による議論の統合がCronbachやRossiなどの理論家によって試みられている。それは、単純な利用ではなく、「より洗練された道具的利用 (more sophisticated instrumental use)」と「啓発的利用」との統合の試みである。彼らが論ずる「より洗練された道具的利用」とは、利害関係者が、評価によって答えられるべき質問を設定し評価の進捗を監視することを内容とする。具体的には、評価者が潜在的利用者を評価前に特定し、潜在的利用者が持つ「道具的利用」についての関心を発見し、評価中も頻繁にそれらの人々との接触を図る。また、利害関係者が操作可能な事項に焦点を当て、情報の質を損なうとしても時宜を得た方法でこれら全ての手続を行う場合に、「道具的利用」が支持されると考えた。また、CronbachとRossiは、「啓発的利用」についても重視した。彼らはWeissと同様、社会問題の根本的要因を扱う質問、および社会問題を変革しようとする努力が困難に直面している理由について評価者に疑問を投げかけている。また、「道具的利用」の評価に対して「啓発的利用」の評価は、迅速な情報の提供又は利用へのプレッシャーが存在しない場合に実施するのがふさわしいことから、アカデミックの評価者または潤沢な資金を持つシンクタンクによる実施がふさわしいと指摘している。

予算と評価のリンクについては次のように論じられている。評価結果はプログラム予算に影響を与えることもあるけれども、政治の複雑性に鑑みて、予算の決定を評価からの発見のみに基づくことは困難なことから、評価は短期的にはプログラム予算に顕著な影響を与えることはほとんどない。むしろ、評価は、プログラムの優先順位を変えることなど予算レベルにおける変更よりも政治的要素が少ない事項に対して影響を与えることができる。また、迅速な「道具的利用」については、「道具的利用」に関する事例の報告が少ない現在では、かつてほど重要と考えられなくなってきている。

これらの議論を経てなお未解決の重要事項として次の3点が指摘されている。第一に、「概念的利用」または「道具的利用」

のどちらが優先的に扱われるべきか。第二に、評価者は意図された評価利用者を特定し、それらの人々の要求に注意を払うべきか。そして、第三に、評価の利用を促すものは何かである。

(3) 日本の政策評価制度の分析

上述のとおり、「単純な道具的利用」から、「より洗練された道具的利用」と「概念的利用」の起こり、そしてその両者の統合へ、という流れで、評価研究における利用をめぐる議論は展開してきたといえる。現在においても、「道具的利用」と「概念的利用」をどういった状況においてどちらを優先的に利用すべきかについては、解答が出されていない。

日本の政策評価制度においては、事業評価（事前）に判断情報としての機能を、実績評価に成績点検情報としての機能を、そして、総合評価・事業評価（事後）に問題解決情報としての機能を期待している。実際には、第5章に見たとおり、予算編成および事業採択の判断を目的とした評価情報の利用が進められ、事業評価（事前）と実績評価が先行している。また、総務省から公表される各年度の政策評価反映状況の存在から考えても、政策評価の結果が迅速かつ直接的に利用されることが期待されているといえる。よって、政策評価の結果は基本的には「道具的利用」が想定されているとすることができる。しかし、評価研究においては、その「単純な道具的利用」は困難であることから、「より洗練された道具的利用」へ、つまり、評価者を通じて利害関係者が評価の視点を設定し、評価の進捗を監視することを内容とするものに変容している。また、「単純な道具的利用」への疑問から生じた「概念的利用」は、社会問題の説明および代替案の提示を行うものであるけれども、政策評価制度において当然のものとは考えられていない。もしありうるとするならば、問題の特定と説明、そしてその問題解決のための代替案の提示を行う場合に、総合評価が「概念的利用」に該当するということもできる。しかしこれさえも、実績評価の延長線につくられたものであったり、社会的問題の説明や代替案の提示を行わないなど、必要な情報が不十分であることが多い。また、「概念的利用」が要求している意思決定に与える影響についても、総合評価は未知数である。

日本の政策評価において、今後暫くの間は、事業の採択と予算の判断に直接的に利用可能な評価への要求が高まっていくと考える。それらが、内部評価としての政策評価の利用へのインセンティブとして有効であると行政機関に認識されているためである。よって、今後さらに、「道具的利用」が志向されるだろう。但し、この「道具的利用」は、社会学者による真実の追究から生じた情報は利用されるはずであると考え、「単純な道具的利用」とも、また、広く利害関係者を巻き込んだ「より洗練された道具的利用」とも異なる、行政機関の職員による行政機関のための評価という、日本の政策評価制度に独自の「道具的利用」が続くと予想される。よって、行政機関にすでに存在する価値観や仕組みを反映した評価が今後も続いていく可能性は高い。しかしこの点について、先行研究にあるとおり、予算と評価のリンクに関しては、外部要因を特定する作業を伴う必要があり、また、政治的要素の多い予算決定に対して評価結果の利用は限定的にならざるをえないことなどから、評価結果の「道具的利用」に過大な期待を持つことがあれば、早くに失望を迎えるかもしれない。一方、個別評価において利害関係者の参画を評価手法に含めたり、各省のパブリック・コメントなどの制度を利用した、行政機関に対する多様な利害関係者による積極的な意見表明が進展した場合には既存の価値観や意思決定に変化が生じる可能性も考えられる。

6.3.2. 予算情報としての評価の利用（省略）

6.3.3. 評価によって産出された情報が消滅する危険性に対する回避策（省略）

<本抜粋における引用文献>

- Hatry, Harry P. 1999. *Performance measurement getting results*. Washington DC: Urban Institute Press.
- Haveman, Robert H. 1987. "Policy analysis and evaluation research after twenty years." *Policy Studies Journal*, Vol. 16, No. 2, Winter, pp. 191-218.
- Love, Arnold J. 2001. "The future of evaluation: Catching rocks with cauldrons." *American Journal of Evaluation*, Vol. 22, No. 3, Autumn, pp. 437-444.
- Owen, John M., and Patricia J. Rogers. 1999. *Program Evaluation*. California: Sage Publications.
- Rossi, Peter H., Howard E. Freeman, and Mark W. Lipsey. 1999. *Evaluation a systematic approach sixth edition*. California: Sage Publications.
- Shadish, William R., Jr., Tomas D. Cook, and Laura C. Leviton. 1991. *Foundation of Program Evaluation*. California: Sage Publications.
- Smith, M. F. 2001. "Evaluation: Preview of the future #2." *American Journal of Evaluation*, Vol. 22, No. 3, Autumn, pp. 281-300.
- Stufflebeam, Daniel L. 2001. "Evaluation Models." *New Directions for Evaluation*, No. 89, Spring, pp. 7-98.
- Vedung, Evert. 2000. *Public policy and program evaluation*. New Jersey: Transaction Publishers.