

政策評価研究の黄昏

メタデータ	言語: jpn 出版者: 明治大学専門職大学院ガバナンス研究科 公開日: 2023-05-31 キーワード (Ja): キーワード (En): 作成者: 西出, 順郎 メールアドレス: 所属:
URL	http://hdl.handle.net/10291/00022950

政策評価研究の黄昏

西出 順郎

1 はじめに

本稿は、我が国の行政機関が実施する評価に対する研究のこれからを、自らの過去の研究を省みつつ、想察したものである。国・自治体においては、1996年に導入された三重県の事務事業評価システムを端緒に、政策や事業の管理を全組織的に実施する事後的評価の仕組み（システムや制度）が導入された。しかし、この仕組みは未だ政策の改善・見直し等には活かされているとはいいい難く（益田2021, 42; 原田2022, 150）、関係者の間で役に立っているという「実感」は乏しいようである（後2022, 199）。

日本学術会議の声明「科学者の行動規範：改訂版」（平成25（2013）年1月25日）によれば、科学とは「合理と実証を旨として営々と築かれる知識の体系」であり、科学研究とは「人類が未踏の領域に果敢に挑戦して新たな知識を生み出す行為」であり、科学と科学研究は「社会と共に、そして社会のためにある」という。そのうえで科学者に対し「社会に対する説明責任を果たし、科学と社会、そして政策立案・決定者との健全な関係の構築と維持に自覚的に参画する」ことを求めている。概して行政学や公共経営に関する研究では、その多くが現実的な問題に対する解決策を見出すことを目的とした応用的な性格を有している（Van Thiel 2014, 28）。行政機関の評価に関心を抱く研究者が「政策を科学的に検証するのが評価であり、この営為は社会のためになる」という命題を背

負いつつ、「(役立たないという) 実感を発露させる仕組とは何か」「その原因は何か」「それはどう改善されるべきか」といった研究ベクトルを好むのも、いわば自然の成り行きといえる。

一般的に政策もしくは行政評価と称され、現在、国・自治体が実際に取り組んでいる政策や事業の業績評価の仕組(以下、「政策評価」¹⁾)という)の研究(以下、「政策評価研究」という)は、政策評価の導入・普及とともに歩み、四半世紀にわたり蓄積をもたらした。政策評価研究の蓄積は政策評価の実務(以下、「政策評価実務」という)に貢献している筈なのだが、先に触れたように、政策評価実務に対する懐疑的な実感は未だ払しょくされてはいない。近年においては、同実務は「お手盛り評価」「評価疲れ」といった批判を暗黙的な基本前提とみなし、また、同研究はそれらの議論から距離を置きはじめている、かのようにもみえる。

本稿では、政策評価研究のこれまでの軌跡とこれからの軌道を考察しながら、自身の過去の研究事例とともに実証的アプローチの意義について論じてゆく。近年、政策評価研究は「魅力が乏しい研究素材とみなされている」ように思う。本稿が同研究の将来に対し、細やかながらも何らかの一石を投じることができたらと考えている。

2 当初の研究動向：評価技術への関心

行政学・公共政策の研究は、大まかに①公共政策そのものが研究主題となる研究、②特定政策の発展に貢献する研究、③特定の政策課題の解決方策を見出すための研究、の3つに分類することができる(Van Thiel 2014, 5)。課題の解決方策には政策手段(policy instruments)のみならず管理技術(management techniques)も含まれているゆえ、政策評価を組織内政策の手段と見立てれば、

¹⁾ ここでの「政策評価」はあくまで我が国での現行評価の仕組を指し、一般的な意味での政策評価より限定的に用いられている。

政策評価研究は②及び③の範疇に入ろう。よってまずもつての「宿題」は、政策評価の仕組が掲げた目的の達成（すなわち政策評価の実効性）に対する貢献であり、貢献のための解を見出すこととなる。その近道として、同研究は評価結果の質を直接左右する仕組の技術に自らの関心を集中させる。例えば、定量・定性といった分析手法の問題や数値指標の適切性、外部評価の強化等である²⁾。

評価技術への関心は、評価が科学性を要する営為であるせい、評価結果を産出する具体の仕組に学術的な研究作法を要請し、評価の実務がどう考えようと、その技術的な良し悪しを問おうとする。その結果、評価に従事する者の行動様式は自ずと科学志向であるという、暗黙的な前提に陥りやすい。特に、評価という営為に政治的な側面があること（Weiss 1997, 312、訳書：411）を忘れ、評価による政策の質の統制を強く念じている場合である³⁾。

評価の実務も評価の前提である科学性には関心がある。そのために政策評価のプロセスに研究者の参画を求める（Weiss 1997, 35、訳書：45）。利害を有しない評価者が従事することで市民からの信頼を得ることが重要だからである（*ibid.*, 37、訳書：48）。評価作業の内実是不分明であっても、外形的な仕組の問題点を評価の実務者との議論でより認識することができ、研究者にとってもその要請は悪い話ではない。

過去の経験が乏しかった政策評価の導入当初は、両者の意図はともかく、政策評価研究と政策評価実務との距離感は近かった。その相互依存的な関係性が政策評価研究の関心を仕組の技術性に係る議論へと導いていったといえる。

²⁾ 国や自治体の政策評価に関する先行研究については西出（2020）第3章を参照されたい。

³⁾ このような暗黙的な前提があるせいだろうか、例えば国の評価制度では、各府省が「評価結果の政策への反映状況」の報告書を毎年作成・公表している。

3 研究関心の移行：仕組の技術から運用・影響プロセスへ

行政学の研究は、公的セクターの問題や課題を研究し、解決策を見つけ出すことが目的である。しかしながら、典型的な政策指向の研究（policy-oriented research）の結果は必ずしも実務に同意・適用されるものではなく、行政への政策提言はしばしば「お蔵入り」となる（以上、Van Thiel 2014, 5-6）。ある程度の月日が流れると、政策評価研究による助言や批判は、政策評価実務がそれらに何ら意味はないと断じたが如く、同じ道を辿ることになる。政策評価は抜本的な改善がなされることなく、「活かさず殺さず」のような硬直状態にあるかのようだ⁴⁾。先の「実感」に鑑みれば、むしろ政策評価研究による指摘は、仕組の簡便化や廃止といった無難な航路へと政策評価実務の舵を切らせたかのようにもみえる⁵⁾。

そもそも政策評価実務の従事者（以下、「評価従事者」という）には外部からの批判で頭を悩ます必要などない。政策評価の仕組では、評価結果の直接的な活用者は評価従事者であり政策立案実務の従事者でもある。如何に第三者が評価結果の質の粗悪さを批判しても、そのような彼ら/彼女が評価結果には意味があると主張すれば⁶⁾、活用者がそう断じる以上、それには意味があるとまずは認めざるを得ない。また、仮に技術性的の問題を解決しようとする、分析データの確保や人材育成に莫大な予算投入が求められる。にもかかわらず自らが科学性を追求したせいで、自らの政策効果が否定されることになったら、それはまるでミイラ取りのミイラ化である。となると、政策評価研究の技術批判は非

⁴⁾ 例えば、自治体の場合については西出（2016）を参照されたい。

⁵⁾ 例えば、2014年3月に総務省が公表した「地方公共団体における行政評価の取組状況等に関する調査結果」では、52団体が政策評価を廃止していたことが明らかになっている。国の制度の簡素化については西出（2020）第10章を参照されたい。

⁶⁾ 各府省が毎年公表している「評価結果の政策への反映状況」報告書のとおり、評価結果は次年度政策に反映されていることになっている。

現実的で所詮対岸の火事の話と、政策評価実務は同研究と一線を画すようになる。同実務による同研究に対する深層的軽視は自然の成り行きなのである。

聞く耳を持たすには、政策評価研究は対岸へと赴き、政策評価実務をつぶさに観察し、それが抜本的な改善に取り組まざるを得ないような証拠を見つけなければならない。硬直化の原因が失火なのか放火なのか、その追究のため、研究関心の一部は仕組の技術的な議論から運用・影響に関する説明へと移行していったのである。もっとも政策評価の実証研究は導入当初から既に取り組まれており⁷⁾、その成果が表層的な関係になりつつあった政策評価研究と同実務の関係に同様に刺激を与え、同研究の軌道の変化をもたらしたともいえよう。

4 実証的アプローチの事例

実証的に政策評価の作業を考察すれば、評価技術の解釈だけではみえない、「宿題」の解がみえてくる。以下は、その考察事例として西出（2020）を取り上げる。もちろん、これが政策評価研究における実証的アプローチの良い事例の一つというわけでは決してなく、研究方法上の限界や課題は少なくない。よってここでは「宿題」の解の一例と同アプローチの難しさを主に述べることとなる。

(1) 目的

本考察の目的は、既往研究から国の政策評価制度（以下、「評価制度」という）の実効性が脆弱であるとの前提に立ち、その要因を「評価従事者の評価行動が評価結果を歪めるありさま」から説明することにあつた。ここでいう「評価行動」とは、政策評価を実施する行政機関内の評価作業に係る一連の動きを、また、「評価結果を歪める」とは、制度趣旨に沿った厳格で客観的な評価ではなく、同趣旨とは乖離した目的をもって（評価従事者が）評価結果を意図的に産出することを指す。同従事者の何らかの意図と評価結果との因果的説明を一般化で

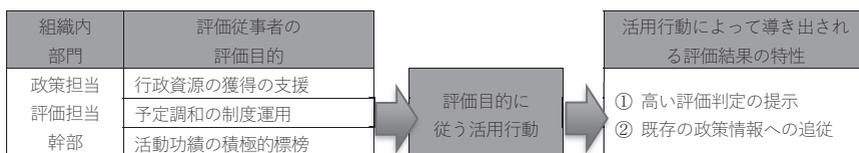
⁷⁾ 例えば、田邊（2006）、原田（2011, 2015）などが挙げられる。

きるのであれば、政策評価の実効性について従来とは異なる論点を提供できる、という想定の下、本考察に取り組んだ。

(2) 考察の枠組み

本考察では、評価従事者の評価目的と評価結果の何らかの因果的なつながりを解明すべく、評価行動の仮説を導き出し、同仮説の実証的考察を試みたものだが、その設定にあたっては、公的な組織構成員の評価の取組という点に着目した。具体的には、官僚制研究及び公的活動の評価理論の各知見を基に、評価行動のありさまが演繹的に推論されている。ここではそれを「作為的評価行動」と名付けた（図表1参照）。

図表1 作為的評価行動の仮説⁸⁾



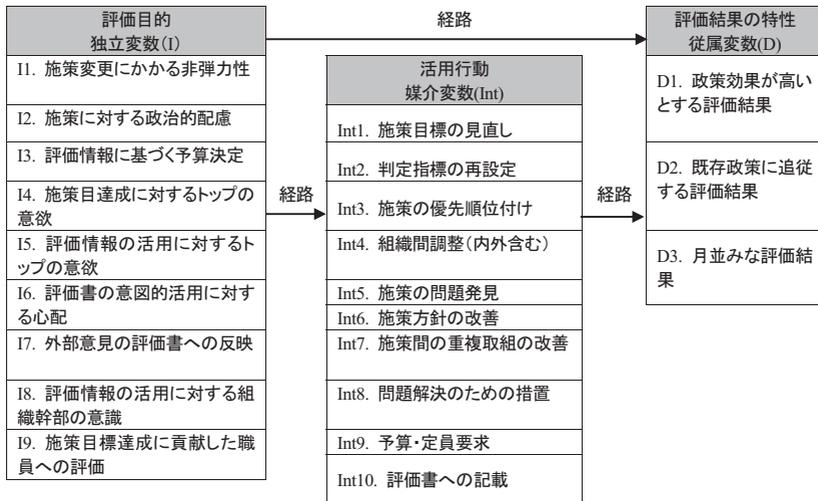
出所：西出（2020）を一部修正

作為的評価行動の仮説を検証するため、本考察では、①少なくとも1つは仮説に従う事実が存在することを確認する、②（確認できれば）仮説を構成する有意な独立変数を抽出し、モデル化する、③（モデル化できれば）その因果的連鎖を説明する、という3つの分析手順を踏んだ。ここでの想定論理は、仮説は少なくとも実在し、そのモデル化及びその因果的メカニズムの説明が可能であるなら、同仮説を一般的に説明できるということである。分析方法としては、①政府資料に基づく実績評価事例の内容分析、②評価実施機関の課長職へのアンケート調査結果に基づく量的分析、③評価部門へのインタビュー調査結果に

⁸⁾ なお本仮説は、事例分析の結果、「中庸化された情報の提示」という評価結果の特性を付加する修正がなされた。

基づく質的分析を用いている。上記①では、文部科学省の法科大学院制度に係る実績評価事例（同制度開始翌年度の平成16（2004）年度から平成20（2008）年度まで計5年間の実績評価の結果）から、同仮説の発現の有無を探索した。②では、同仮説を操作化したうえで、評価実施機関の課長職の全職員（計983人）に対してアンケート調査を実施（有効回収数（率）160（16.3%））し、当該調査結果を階層的重回帰分析及びパス解析にかけ、同仮説のモデル化を試みた。後述する作為的評価行動メカニズムの各独立及び媒介変数⁹⁾は、あらかじめ設定された同行動の想定経路（図表2参照）から両分析結果に従い抽出されたものである（5%水準）。そして③では、評価実施機関すべて（計21機関）の評価部門職員を対象に同モデルを提示した半構造的インタビューを実施し、同モデルへの因果的説明の付与を試みた。

図表2 作為的評価行動の想定経路



出所：西出（2020）を一部修正

⁹⁾ 当該想定経路を構成する各変数は、GAO（米国政府監査院）の連邦職員向けの調査項目（GAO2013）から、当該仮説との関係性に鑑み、選択されたものである。

(3) 考察の結果

各分析の結果を踏まえ、評価制度における3つの作為的評価行動のメカニズムが提示された(図表3参照)。1つめは評価従事者が行政資源の獲得を支援もしくは活動業績を標榜するために評価制度の利用を目論み、高い評価判定を評価結果として提示する評価行動である。2つめは評価従事者が行政資源の獲得支援もしくは予定調和的に着実で円滑な制度運用のために評価制度の利用を目論み、評価結果が既存の政策情報に追従する評価行動である。最後は評価従事者が行政資源の獲得支援もしくは予定調和的に着実で円滑な制度運用の遂行のために評価制度の利用を目論み、月並みな評価結果を提示する評価行動である。これらの可視化によって、本考察は「評価従事者の評価行動が評価結果を歪める構図」を因果的に描き出したことになる¹⁰⁾。

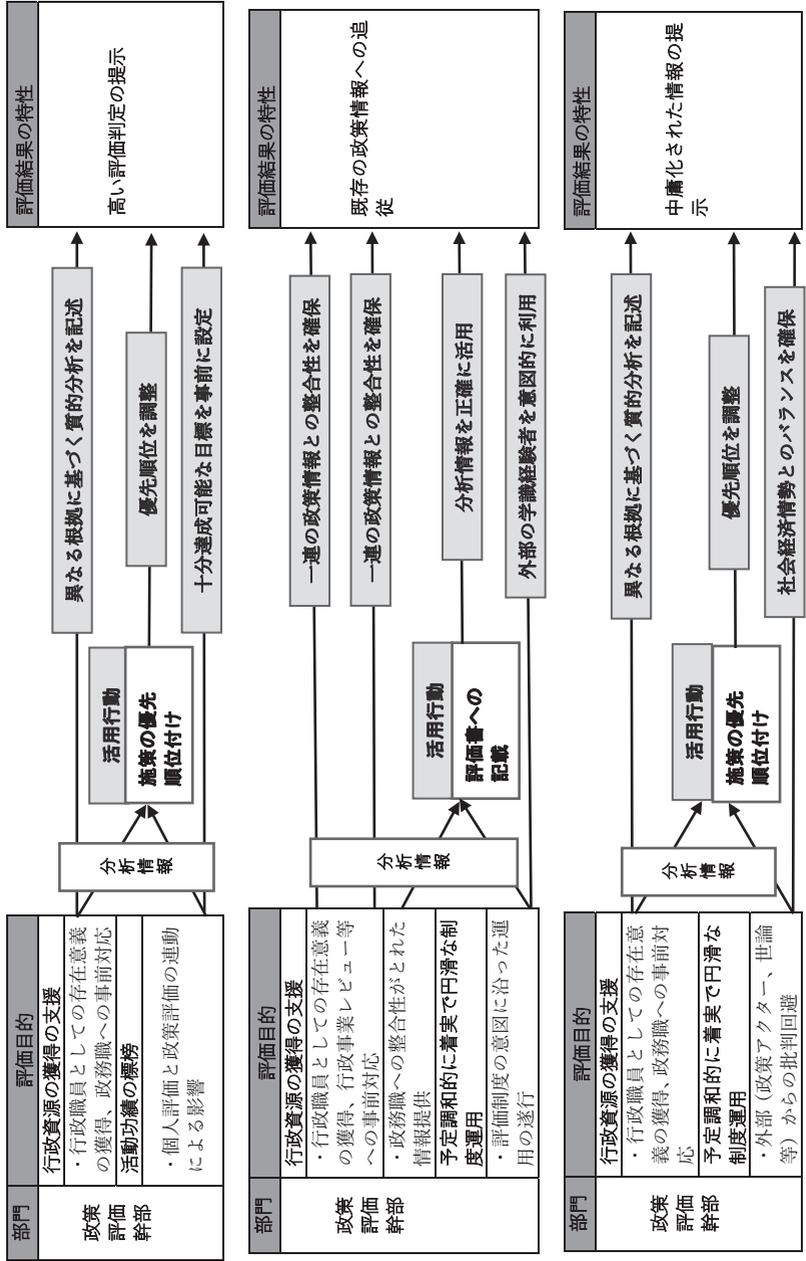
(4) 分析の限界

但し本考察の結果は、各モデルの決定係数や適合度¹¹⁾に鑑みると、必ずしも高い説明力を有しているわけではなく、他の説明できる評価行動や評価結果に影響を及ぼす非科学的要因の探索を自らの課題として提示してくれる。本結果がまだまだ発展途上にあるという、一つの証左である。このように政策評価研究の「宿題」を解くには、国の公表資料だけでなく、分析データとして行政職員の認識にも頼らざるを得ない。しかし、そのようなデータにおいて結果の信頼性及び妥当性、再現可能性を安定させるのは決して容易ではない。実証的考察の質を高めようとすればするほど、これらの問題と多々向かい合うことになる。もし「お手盛り評価」が評価従事者側の共通認識とされるのであるなら、例え

¹⁰⁾ なお研究全体としては、実証的考察の補完的作業として政策評価制度の立案及び実施過程を探索的に考察し、作為的評価行動を生じさせているであろう要因について制度及びその運用という視点からも論じている。

¹¹⁾ 「高い評価判定の提示」「既存の政策情報への追従」「中庸化された情報の提示」の各モデルの決定係数及び適合度(RMSEA)については、.233と.052、.421と.000、.374と.000、という結果となっている。

図表3 作為的評価行動のメカニズム



出所：西出 (2020)

ば、政策評価に関する調査への協力、すなわちアンケート調査の回収率の向上を得る工夫などそう簡単には見つからない。また、母集団が1,000足らずであるため、分析に用いる変数数も自ずと慎重になり、その数が仮説構築を検討する際の制約的条件にもなる。また、類似の調査を実施するとすると、アンケートもインタビューも対応する職員はほとんど異なる。さらには、前回調査の結果が公表されている以上、内的妥当性への脅威への対応はより難しくなる。チェリー・ピッキングやテストング効果のようなバイアスもより強くなる。調査への協力自体、得られにくくなるかもしれない。仮に同じ調査を今再び実施するのであれば、分析結果の再現性についてはかなりの不安感を抱かざるを得ない。

5 政策評価研究の行く末

(1) 実証的アプローチの意義

政策評価の作業の内実を説明できれば、いわゆる「お手盛り批判」を受けながらも、なぜその仕組が抜本的に変わらないのか、なぜ実効性が乏しいとされるのか、これらの解に着実に接近できるであろう。また、このような研究成果は従来その裏書的な役割を担うことにもなろう。ただその内実を説得的に論じることは、評価作業の動きを直接観察できず、また主要な分析データを評価従事者の認識に依存するがゆえにかなり難しい。本来なら、やはり意図や自己認識などの抽象的で観察できない概念は分析に用いるべきではない (King, Keohane and Verba 1994, 109-110、訳書：131-132) のであろう。しかし「もっと危険なのは探しているものが暗闇にあり努力すれば見つかるにもかかわらずむやみやたらに「明りの下を探すこと」である」(Van Evera 1997, 47-48、訳書：49) なら、また、彼ら/彼女の認識の奥底に「宿題」の解が横たわっているのであるなら、それを探る努力を簡単に諦めるわけにもいかない。確かに、政策評価の仕組に執着する実証的アプローチは分析結果の妥当性や信頼性が安定的ではないかもしれない。しかし同アプローチは、評価の仕組が如何に合目的で

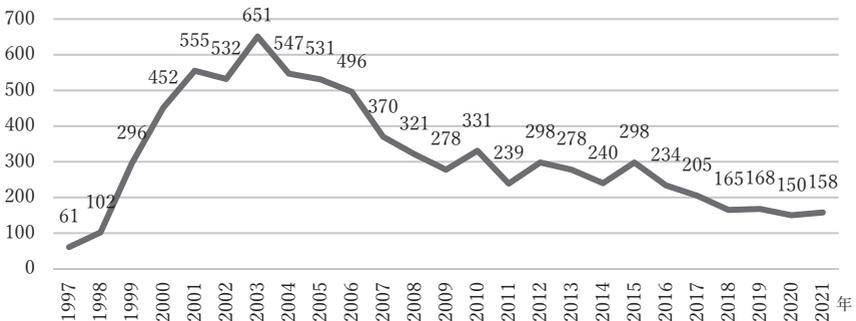
あってもその機能上の脆弱性を明示できる、政策評価への直接的貢献に資する説得的な研究デザイン上のベクトル（以下、「研究ベクトル」という）なのである。

（2）2つの研究ベクトル

「日本の評価システムはガラパゴス化（国際標準から逸脱）した。実務が評価学の知見を十分いかしきれていない」（山谷 2021, 159）。政策評価研究と同実務は細長いテーブルの両端に座り、これまで議論を続けてきたようだ。図表4は、1997年から2021年の25年間において、「政策評価」もしくは「行政評価」がタイトルの中に付された、書籍・雑誌記事・論文総数の推移を示したものである。

政策評価に係る研究論文に限定されたわけではないが^{s12)}、その数は同評価が全国的に普及した2000年代初め以降、減少傾向が続いている。政策評価実務のダイナミズムを湧出できない政策評価研究への諦観からきているのかもしれない。

図表4 政策評価に関する文献数（書籍及び雑誌記事・論文）



（出所：国会図書館 DB 検索から筆者作成）

¹²⁾ 当該検索においては、政策評価研究のみならず個別政策の評価研究や一部の自治体の評価報告書等もカウントされている。

斜陽化ともみてとれる政策評価研究をよそに、近年、EBPM や独立財政検証機関、さらにはアジャイル指向の政策形成・評価等、政策評価に近しい実務の動きが研究素材として注目されている。もし同研究における実証的アプローチが進展しないようであれば、同研究は、政策評価実務に貢献できず、自ずと終焉を迎えてしまうのであろうか。これまでの軌跡から同研究の軌道を見据えようとするなら、今後の同研究には次のような2つの研究ベクトルが浮かび上がってくる。1つ目は様々な関心から研究設問を見出し、同研究の「宿題」とは距離を置き続ける研究ベクトルである。市民参加による評価や人材育成のための評価など、実際の政策評価と観念上の評価枠組みとの境界領域で現実的なような空想的なような処方箋を提示する、探索型のアプローチを指向するものである。2つ目は上述のような、我が国で装置化されていない、政策を評価・検証する取組に関する研究ベクトルである。これは、海外の事例等を紹介もしくは比較考察しそれらの推進・普及を目指す、唱導型のアプローチを指向している。前者は行政機関における評価の研究のすそ野を拡大し、政策を評価する意味をより広く解釈させることで、後者は我が国の政策の評価・検証の枠組みの再検討を促すことで研究に対する貢献を果たすことになるであろう。その一方で従来の政策評価研究は傍流となり、同研究への「宿題」は不問に付されることになるだろう。さらに、前者の場合は政策評価への言及自体がより疎かになり、後者の場合、同研究上での議論が別の土俵の上で繰り返されることになるかもしれない。政策評価研究者にとってそれはまさにデジャブとなる。

科学には知見の蓄積が求められる (Van Thiel 2014, 26)。いずれの終着地においても、崩れ落ちる瓦礫を積み戻すかのような政策評価研究の黄昏が眼前に広がっているかのようである。

6 結びにかえて

本稿では、政策評価研究の軌跡と軌道を考察し、また自身の実証的考察の限界等に触れながら、同研究における実証的アプローチの意義について論じてき

た。政策評価研究に課された「宿題」を解くことに自らの意義を見出すのであるなら、同研究のこれまでとこれからの鑑みても実証的アプローチが政策評価実務を得心させ、適用できる成果を出してゆく有用な研究ベクトルであるというのが、些か規範もしくは情緒的かもしれないが、ここでの帰結とするところである。

近年、政策評価と近い取組の文献に目を通すとき、かつて既往研究をレビューしていた際に遭遇したデジャブを思い出すことがある。1950-60年代の行政測定、70-80年代のPPBS、プログラム評価やマネジメント・レビュー、90年代の政策評価、近年のEBPM等々、いずれの研究においても指標の数値化や分析手法、外部要因の統制といった諸課題が繰り返し議論の俎上に載せられている。80年代、プログラム評価の技法を重視する研究者は業績測定でしかない政策評価に抵抗感を覚え、2000年代初めにロジックモデルの政策評価への適用を論じた研究者はEBPMでのロジックモデルの議論に多少なりとも失望しているのではあるまいか。政策評価の「活かさず殺さず」は政策評価研究の「活かさず殺さず」と対をなす。同研究の説得力や断定力の弱さに課題があるのなら、また、行政学研究の目的が公的セクターの課題解決の発見にあるのなら、政策評価の解剖には実証的アプローチのより一層の充実が求められる筈である。

政策評価研究は黄昏時に漂い続けるわけにはいかない。同研究は自らが背負った「宿題」の解を総括することができるのか。政治が「情熱と判断力の2つを駆使しながら、堅い板に力をこめて、じわっじわっと穴をくり貫いていく作業」(ヴェーバー 1980, 105) というのなら、同研究も評価に潜む政治と対峙する研究として、自らの「見当識」を高め続けてゆかねばならぬであろう。

(付記)

本稿は、2022年度日本行政学会研究会分科会 A2「政策評価の実施・影響プロセスの実証研究」に係る報告論文を一部修正したものである。

(参考文献)

- 後房雄 (2022) 『地方自治における政治の復権：政治学的地方自治論』 北大路書房
- 田邊國昭 (2006) 「政策評価制度の運用実態とその影響」『レヴァイアサン』、No. 38、86-109
- 西出順郎 (2016) 「自治体評価を振り返る：「活かさず殺さず」の20年」『日本評価研究』、Vol.16、No.1、17-30
- (2020) 『政策はなぜ検証できないのか：政策評価制度の研究』 勁草書房
- 原田久 (2011) 「各府省における実績評価：「政治主導」は10年の制度運用を乗り越えうるか？」『評価クォーターリー』、No.18、2-12
- (2015) 「府省における評価連携」『評価クォーターリー』、No.34、2-9
- (2022) 『行政学第2版』 法律文化社
- 益田直子 (2021) 「政策評価における負の側面の抑制について」『評価クォーターリー』、No.57、34-46
- 山谷清志 (2021) 「政策学における評価理論の貢献：日本の評価システムから」『同志社政策科学研究』、Vo. 22、No. 2、159-172
- ウェーバー、マックス (1980) [脇圭平訳] 『職業としての政治』 岩波文庫
- King, G., Keohane, R. O. and Verba, S. (1994) *Designing Social Inquiry: Scientific Inference in Qualitative Research*, Princeton University Press、真淵勝監訳 『社会科学のリサーチデザイン：定性的研究における科学的推論』 (勁草書房・2004)
- Van Evera, S. (1997) *Guide to Methods for Students of Political Science*, Cornell University Press、野口和彦他訳 『政治学のリサーチ・メソッド』 (勁草書房・2009)
- Van Thiel, S. (2014) *Research Methods in Public Administration and Public Management: Introduction*, Routledge
- Weiss, C. H. (1998) *Evaluation*, 2nd edition, Prentice Hall、佐々木亮監修 『入門

評価学：政策・プログラム研究の方法』（日本評論社・2014）

（資料）

総務省（2014）「地方公共団体における行政評価の取組状況等に関する調査結果」

総務省ホームページ（2022年4月7日閲覧）https://www.soumu.go.jp/main_content/000278817.pdf

GAO（2013）*Survey of Organizational Performance and Management Issues*,

Government Accountability Office（2022年4月7日閲覧）<https://files.gao.gov/special.pubs/gao-13-519sp/results.htm>