

評価システムとガバナンス

日本の ODA における評価ポリシー

三上 真嗣

概要

評価の氾濫が懸念されている。理論は次々と提唱され、時間の流れとともに実務は変容して複雑化する。評価実践の複雑な全体像は、評価の実務者や専門家以外には説明できない、よくわからないものになる。しかし、評価の氾濫は本当にガバナンス機能を阻害しているのだろうか。この解明には、複雑化が進む評価実践の実態を理論的に把握する方法を確立する必要がある。その具体的な方法として「評価システム (Evaluation System)」という新しい枠組みに注目する。

本稿では次の3点を検討した。第1に、先行研究を整理して評価システムの枠組みを提示した。第2に、評価システムの理論的類型をガバナンス機能と関係させて導出した。第3に、日本の評価システムの一例として、ODAの行政活動における政策評価の複雑な実態を評価システムの枠組みを通じて考察した。具体的には、評価ポリシー (evaluation policy) と評価情報の2点に焦点を当てた。

各々の評価システムは相互利用されることで、重層で複雑な評価システムを全体として形成していた。そこでは、管理志向型の結果がアカウントビリティ志向型によって利用される流れが見られて、ガバナンス機能が確保されていた。評価の氾濫がかならずしもガバナンス機能を損ねるわけではない一例が示された。同時に、評価の氾濫に対する処方箋として評価システムのガバナンスが必要であるという課題が提示された。

1. はじめに

実務においては、社会の要求に応じて評価実践が増加し、絡み合っって複雑な様相を示している。次々に提唱される評価理論も評価実践に取り込まれていくが、時間の経過とともに組織の生理や病理の影響を受けて変容して複雑化が進む。そのため、シンプルな理論からは乖離していく。結果として評価実践の複雑な全体像は評価の実務者や専門家以外には説明できない、よくわからないものになって評価の氾濫が懸念されてしまう。しかし、この複雑な実態を解明しないままに評価に対する失望を結論づけるのは時期尚早である。評価の氾濫は本当にガバナンス機能を阻害しているのだろうか。

この疑問を確かめるためには、実務における複雑な実態を理論的に把握する方法を確立する必要がある。具体的には、「評価システム (Evaluation System)」という新しい枠組みを用いたアプローチに注目する。これまでは評価の制度を分けて個別に説明しがちであったが、より複雑に入り組んだ評価実践を体系的に把握することができる。

本稿では次の3点を目的にする。第1に、評価システムの枠組みについて整理と考察を行って、日本で培われてきた学問の延長線上に据える。第2に、評価システムの理論的類型を導出する。第3に、日本の評価システムの一例を説明してガバナンス機能が確保できているかの考察を試みる。

具体的には、次のような構成で議論を進める。まず、評価システムに関する先行研究の整理をして、日本の行政学、政策学、評価学に沿った定義を確認する。次に、行政学とガバナンス論

の知見を用いて、評価システムの理論的類型の枠組みを導出する。さらに、実際の日本の政府開発援助（Official Development Assistance 以下、ODA）の行政活動における政策評価の複雑な現状を評価システムの枠組みを用いて描出し、ガバナンス機能がどのように確保されているのかを考察する。なお、ODAを事例に取り上げる理由もその際に述べる。最後に、議論の総括と限界の確認を行ったうえで評価システムのガバナンスを中心に課題と今後の展望を述べる。

2. 評価システム

新しい概念の議論に立ち入る前に、政策評価に関する基本的な定義について確認しておく必要がある。山谷は評価（evaluation）を「何らかの判断や決断をくださる場（人）に、情報を提供する活動」と定義する（山谷 2006：71）。とりわけ政策評価は、政策を評価対象にした情報提供活動全般を指す包括的な概念である。この広義の政策評価と法律で規定される「政策評価」は異なり、「政策評価」は「行政機関が行う政策の評価に関する法律」に規定される制度である。本稿では、かぎ括弧を用いて両者を区別する。

評価理論の提唱は多元化と多様化を遂げた。しかし、実験的に一度だけ実施されたまま組織に完全には定着しないものも多い。理論的に提唱された評価は組織実務の複雑な事情を考慮せず、論理の内部で自己完結的である場合もある。こうした理論も評価の専門家以外にとってはよくわからないので、実務の担当者も配属が変わるたびに評価理論の流行り廃りに振り回されてしまう。

評価実践は、こうした評価理論の流行り廃りの歴史的産物であるともいえる。評価実践における評価理論のこうした流行り廃りは「4つの評価の波と堆積物の共存した集積すなわち、科学主義の波、対話の波、新自由主義の波、エビデンスの波」と表現されることもある（Vedung 2010：265）。この「堆積物の共存した集積」という表現が重要な示唆を与える。すなわち、実務においてはなにか1つの評価理論を選択して用いるのではなく、評価理論の波が寄せては引くなかでその都度、組織に少しずつ蓄積して複

雑な評価実践をつくり上げることを示唆する。その結果、1つの評価理論では説明が困難になる。

この複雑な評価実践を把握するための枠組みが評価システムである。実際に、北ヨーロッパの研究者たちを中心に、複数の研究者たちがこの枠組みを用いて評価に関する研究を行っている（Hanberger 2011；Höjlund 2014；Martinaitis, Christenko and Kraučiušienė 2018；Raimondo 2018）。

2002年に出版された“International Atlas of Evaluation”では日本を含む21の国と3つの機関を対象にした国際比較が行われた（Furubo, Rist and Sandahl 2002）。しかし、それぞれの国の評価は各々が特徴的であったので、厳密な意味での比較は困難であった。2000年代前半までにはOECD加盟国をはじめとしたほとんどすべての先進国で政策評価は広まったのだが、同時に評価をめぐる議論や実践も多元化して多様化したのである。その結果、比較を行ううえで困難が生じてしまう。

この国際比較が契機となって“From Studies to Stream”が著されたことで、評価システムの議論の基礎がつくられる（Rist and Stame 2006）。すなわち、単一の評価を対象にした研究が生み出す知見は、他では応用が効かないという問題が認識された。「評価のシステム（Evaluative System）」が副題であるように、単一の評価ではなく流れ（stream）としての体系的な評価の実態を研究すべきであると主張された。

この「評価のシステム」の影響を受けて「評価システム（Evaluation System）」が新たに提唱された。具体的には、次の4つの特徴を満たすものが評価システムであると定義されている（Leeuw and Furubo 2008：159-60）。

第1の「特有の認識論的な側面（A Distinctive Epistemological Perspective）」は、組織の構成員が他とは異なった特有の評価を行っていることと認識していることである。具体的には、テキスト、ガイドラインあるいはウェブサイトといった評価に対する共通理解を生み出すものによって把握できる。評価を通じてどのような知識を生み出すかもこの側面にかかわる。

第2の「組織的な責任（Organizational Responsibility）」は、組織によって評価が行われて

いることであり、1人の評価者だけによって評価が行われている状態ではないことである。そのためには、少なくとも1つ以上の活動中の組織が存在することが必要である。ようするに、評価に携わる組織が存在しない場合は評価システムたりえない。

第3の「永続性 (Permanence)」は、現在進行系で評価が行われている、あるいは行う予定があり今後も継続することである。この永続性は、定期的な刊行物の有無やマスメディアとの関係、あるいは適切で専門的な学会とのつながりの有無によって把握できる。実験的に導入した一度きりの評価は評価システムではない。

第4の「評価の意図された利用への焦点 (A Focus on the Intended Use of Evaluations)」は、評価を通じて産出された情報が政策決定過程や政策実施過程といった政策過程に対してフィードバックされるように意図されていることである。そのため、評価の利用 (evaluation use) の議論も関係する。

これらの要素を次のように考察すると、評価システムは評価学、行政学および政策学の3つの学問分野の関心が重なったところにある概念であるといえる。まず、評価活動が実際にどのような評価であるかに関心がある点で評価学的な概念である。すなわち、どのような評価が行われているのか、どこまでが評価といえる活動であるかなど評価学におけるこれまでの議論が関係する。

次に、官僚制組織における評価活動の生理と病理に関心がある点で行政学的な概念である。定義の4要素のうち2つが評価の組織的な側面である点からも、この行政学的な側面を重視していると読み取れる。組織的な評価活動の実施については、官房系組織がどのように活動するかを考える必要がある。また、永続性の観点については、組織において評価がどのようにルーティンワークとして定着して実施されているのかを考える必要がある。「評価は真空地帯で行われぬ。評価システムは、組織および組織が組織文化のあれこれで形作り形作られるやり方のなかに埋め込まれる」と述べる研究者もいる (Raimondo 2018:26)。すなわち、組織活動がルーティン化されて、組織に定着するものが評価システムの焦点である。まさに官僚制組織とその組織文化のなかで組織活動がどのように行われ

るかという行政学特有の着眼点がみられる。

最後に、政策過程に評価活動の結果がどのように関係するかに関心がある点で政策学的な概念である。評価結果のフィードバックは、政策過程における政策改善や政策管理の議論にも関係する。評価が産出するものは情報であり、どのような流れが官僚制組織の内外に及んでいるのかを考察する必要がある。また本稿でも用いるように、政策過程に対する間接的な影響すなわち、評価が他の評価の基礎的な情報を提供する場合も考えられる。

ようするに、評価システムは突拍子もない思いつきではなく、各分野で個別に行われてきた議論を自覚的に統合して制度を把握するための枠組みである。だが、他の概念と混同されてしまう場合もある。たとえば、評価制度との区別がつかないおそれもある。また、評価パラダイムや評価トレンドとの違いがあるのかという疑問も呈されている (Palfrey, Thomas and Phillips 2012:37-8)。

評価システムはその他の概念とは関心もアプローチも異なる。たとえば、新制度論は制度を広くルールとして定義することもあり、評価システムは制度の一種といえる。だが、その包含関係だけでは2つの概念が代替可能であるとはいえない。ルールの動作や変容、機能、結びつきといった動態に対する関心を強調する点でシステムの語を用いるほうが適当だからである。しかも、評価制度という語を用いたときに法制度上の評価活動が想起されやすいという懸念もある。

評価パラダイムや評価トレンドとの関係を考えれば、それらに影響された評価実践が経年的あるいは歴史的に堆積して絡み合った実態が評価システム独自の焦点なのである。なお、これまでの議論からもわかるようにシステムの語は、工学的な意味を強調しない。むしろ行政システムの種類であるという側面を強調している。

それでは、評価システムを用いてどのように考察すればよいだろうか。先述した4つの特徴を抽出するために用いることのできる情報源として次の3つが考えられる。すなわち、①評価ガイドラインや大綱などの評価システムを組織的に方向づける評価ポリシー (evaluation policy)、②年次評価報告書などの評価の組織的

な産出情報、③組織文化や慣習などのその他の実務上のルール、である。ただし、③を用いるためには実務に入り込んだり実務者にインタビューを行ったりする必要がある。その準備段階として本稿では①と②を用いて考察を進める。

これらの資料をもとにして、どのように考察すればよいだろうか。先行研究では評価システムの類型化によるアプローチが試みられている (Leeuw and Furubo 2008 : 161 ; Højlund 2014 : 432 ; Martinaitis, Christenko and Kraučiušienė 2018 : 5)。しかし、これらの理論的類型は、類型化の軸としてガバナンスを扱っていない。そこで、ガバナンス機能の確保に対して評価システムが果たしている機能の視点から理論的類型を導出する。

3. ガバナンス機能のための理論的類型

評価システムの理論的類型の枠組みには、次のような先行研究がある。たとえば、「パフォーマンス・モニタリング」、「パフォーマンスの監査、監察および監視」、「(準) 実験的評価とエビデンスに基づく政策の動向」、「アクリディテーションと評価」、「モニタリングと評価」のシステムの5類型が考えられている (Leeuw and Furubo 2008 : 161)。ほかにも、評価の利用 (evaluation use) から分類するものなど、さまざまな評価システムの類型化が試みられている (Højlund 2014 : 432 ; Martinaitis, Christenko and Kraučiušienė 2018 : 5)。しかし、ガバナンス機能との関係による類型はいまだ提案されていない。そこで、本稿では新たにガバナンス機能と関連させて理論的類型を導出する。すなわち、どのように評価がガバナンス機能の確保に貢献しているのかを分類軸にした視角を考える。そのためには、まずガバナンスとはなにかを整理する必要がある。

ガバナンスは多義的な概念であるので既存の類型による把握は困難である。ガバナンスの議論には次の3分野からの定義があると考えられている (Kjær 2004:4)。すなわち第1に、R.A.W. Rhodes に代表される行政学と公共政策のガバナンスである。第2に、J.N. Rosenau に代表される国際関係論におけるグローバル・ガバナ

ンスである。第3に、G. Hyden に代表される比較政治学と開発におけるガバナンスである。

ただし、こうした3つの分野における定義は排反しない。たとえば、ODA における行政活動では3つのガバナンスが併存しながら追求されているだろう。具体的な例として、外務省と独立行政法人の間には行政学と公共政策的なガバナンスが追求され、世界銀行、日本政府、被援助国の政府や住民などの間には国際関係論的なグローバル・ガバナンスが追求される。また、世界銀行が主導して途上国の援助条件としてしばしば掲げられるグッド・ガバナンス (Good Governance) は比較政治学と政府開発援助におけるガバナンスである。

行政学と公共政策におけるガバナンスの論者として分類された R.A.W. Rhodes 自身もまた、ガバナンスを7つに分類している。すなわち、「コーポレート・ガバナンス」、「NPM 型ガバナンス」、「グッド・ガバナンス (good governance)」、「国際的相互依存のガバナンス」、「社会サイバネティック・システム (socio-cybernetic system) としてのガバナンス」、「新しいポリティカル・エコノミー (New Political Economy)」、「ネットワークとしてのガバナンス」の7つをあげている (Rhodes 2000 : 55-63 ; 山谷 2006 : 101-5)。

このように多義的な概念であるガバナンスを1つに定義して言及するのは困難であるが、それでも少なくとも共通する要素はガバナンスの語源でもある舵取り (steer) の意味である。たとえば、「ガバナンスは2つの意味をもつ。すなわち、一方で20世紀末に現れたような国による外部環境に対する適応の経験的表出を意味する。他方でガバナンスは、社会システムと、大半の場合で、その過程における国の役割との調整の概念的あるいは論理的表出をも示す」と言及される (Pierre 2000 : 3)。この言及からも舵取りの意味が時代とともに変化しつつ残り続けていることがわかる。より端的に舵取りに関連して、ガバナンスは「政府より広義のなにかを指し、それは舵取りおよびゲームのルールに関連するということができる」と考えられている (Kjær 2004 : 7)。

この「ルール」は、実際には政府による公共政策に限定される場合が多いだろう。そのような文脈では、「政策を作り実施する、換言すれば社会を舵取る政府の能力 (capacity)」とガバ

ナンスは定義される (Pierre and Peters 2000:1)。その具体的なガバナンスの活動としては、「共通の優先順位の社会に対する明示」、「一貫性 (coherence)」、「舵取り (steering)」、「アカウントビリティ」の4つが考えられている (Pierre and Peters 2005 : 3-5)。「共通の優先順位の社会に対する明示」とは目標の社会的な合意形成であり、ある種の政治である。「一貫性」とは組織間の整合性確保である。そして、「舵取り」とは社会の変化に対する政策手段の適応である。

なお、ガバナンスが舵取りの意味から逸脱しない限り、ガバナンス機能の達成に評価は必要である。なぜなら、舵取りは自他の状況が把握できない限り不可能だからであり、「情報提供のツール」(山谷 2012 : 12)である評価が必要になるからである。評価なくしてガバナンスは成り立たない。

それでは、ガバナンスの活動に評価はどのように関係するのか。先に検討したガバナンスの定義が政策を通じた政府と社会の関係に着目していた点を考えれば、行政における技術の概念を通じて考察できるだろう。西尾勝は、広義の行政、すなわち統治過程における技術を管理技術 (managerial technique) と行政技術 (administrative skill) とに区別した (西尾 1979 : 40)。管理技術とは「官僚制がその第1次的機能である目的活動を効率的かつ能率的に達成するために、その組織態勢そのものの維持発展をはかる第2次的機能」(西尾 1979 : 40)の技術である。他方で、行政技術は、政策立案から政策実現の全過程における技術、すなわち広義の管理技術の範囲外にある、官僚制とその外界にあるコミュニティとの相互作用にかかわる技術である。このことから、2つの技術が用いられる範囲と関係が異なるとわかる。行政技術は行政と社会との相互作用における技術であり、管理技術は官僚制組織の内部で用いられる技術である。

アカウントビリティやマネジメントといった目的をもつ評価はこれらの技術にかかわるツールである。さらに、ネットワーク構造や市場といった代替物によってガバナンスが担われることで官僚制組織が断片化して散逸しつつある状況においては、ガバナンスの態勢そのものの維持発展が要求される。追及されるアカウントビ

リティもまた、市場やネットワーク的なガバナンスの台頭によって変容すると考えられている (Considine and Afzal 2012 : 629)。そこでガバナンス領域とガバナンス領域内部の組織の二重構造が出現したことを考えれば、行政技術と管理技術の間には3種類の技術が見いだせるだろう。すなわち、ガバナンス領域内と外部の社会との相互作用における行政技術、ガバナンス領域内で断片化した組織や社会の間で一貫性を確保する管理的な行政技術、ガバナンス領域内にある組織の態勢の維持発展にかかわる管理技術である。この技術的側面に加えて、その他の情報提供による知的貢献も評価の重要な側面である。なぜなら、専門分野に貢献するための評価研究などは政策評価のルーツのひとつだからである (山谷 1997 : 38)。

以上の議論から、行政の技術の視点をもとにガバナンス活動に関連して4種類の枠組みを導出する(表1)。すなわち、①アカウントビリティ志向型、②一貫性志向型、③管理志向型、④知的貢献志向型の4つの評価システムである。

①アカウントビリティ志向型の評価システムは、行政技術の傾向が強くガバナンス領域内部と外部の社会との間でアカウントビリティを追及する。ガバナンスの活動としては、アカウントビリティ追及、合意形成、舵取りの推進が考えられる。具体的な評価の手法としては、プロセス評価やプログラム評価、エンパワメント評価などがあげられるだろう。

②一貫性志向型の評価システムは、①アカウントビリティ志向型と③管理志向型の中間に位置する。おもに用いられる技術は行政技術的でありつつ、その目的はガバナンス領域それ自体の態勢の維持発展である点が①アカウントビリティ志向型の評価システムとは異なる。ガバナンス領域内にある組織や社会の間で一貫性を確保するためにアカウントビリティが用いられる。すなわち、アカウントビリティの連鎖によるマネジメントが行われている。たとえば、独立行政法人評価やアクリディテーションである大学評価などは、実施機関に対して一定程度の自律的な権限と裁量を認めるかわりに政府の政策体系に沿っているかを問う。ガバナンスの広がりや公的領域に留めおくために用いられるのである。その具体的な手法としては、①と③のいずれのものも考えられる。

③管理志向型の評価システムは、管理技術の傾向が強く組織内部における組織態勢の維持発展にかかわる。具体的な評価の手法としては、業績測定や評定（rating）、政策分析などがあげられる。手段的な船漕ぎ（rowing）のためのマネジメントを重視するので、管理志向型の評価システムしか存在しない場合は、ガバナンス機能の有効性を損ねてしまう。そのため、管理志向型の評価システムがガバナンスにより貢献するためには他の評価システムとの間のフィードバック関係が整備されている必要がある。政策評価の研究では、管理評価の主流化の傾向とその問題点も指摘されている（橋本 2017：32）。

④知的貢献志向型の評価システムは、組織と社会との間でその他の情報提供活動を行う。たとえば、専門分野に対する知的貢献や知的好奇心あるいは科学的興味に基づく調査活動などがある。そのため、アカウントビリティやマネジメントの文脈とは異なる目的として学習（learning）が標榜される場合もある。これらの知見が政策過程にフィードバックされることでガバナンス活動における目標設定を促進する場合もあるだろう。

この4類型を用いてガバナンス機能の有効性について考察していく。具体的には、行政技術的なアカウントビリティ志向型の評価システムに各々の評価システムの産出情報が最終的に結びついているのかに注目する。というのも、ガバナンス活動の多くは行政技術として分類されるので、ガバナンス機能の確保には行政技術的なアカウントビリティ志向型の評価システムが機能している必要があるからである。実際の評

価システムは、1つで複数の機能を果たす場合もあるだろう。その場合は、どの機能が十分であるか不十分であるかを把握する手立てになる。

4. ODA における評価システムとガバナンス

実際の評価システムは、どのように類型化されるだろうか。また、どのように各々の評価システムは結びついているだろうか。この2点を ODA の行政活動における政策評価を対象にして考察する。

ODA の分野を対象にする理由は次の3点である。第1に、透明性が確保されているので信頼性の高い評価関連資料が豊富に公開されており、その入手の容易さから比較の実証的な考察ができるからである。第2に、複雑なアカウントビリティを追及するべく最先端の評価理論が提唱されて、評価実践においても試行錯誤が進む分野だからである。しかも、ODA における評価の理論と実践は国内の他分野にも波及していく傾向にあるので、評価の潮流を把握するのに有用である。第3に、実態の変化に研究が追いついていないので、実務と理論の改善に貢献できるからである。

ODA の行政活動における政策評価は、ODA 評価だけではない。2003 年度末に外務省で行われていた評価として、次の6つの政策評価すなわち「政策評価」、「ODA 評価」、「独立行政法人評価」、「行政評価」、「個別評価」、「外交政策評価パネル」があったと山谷は述べている（山

表 1 行政の技術類型による評価システムの理論的類型

評価システム類型	技術分類 —関係する領域	ガバナンスの活動	評価の目的	評価の例
①アカウントビリティ志向型	行政技術 —ガバナンス領域内と外部の社会間	アカウントビリティ、合意形成、舵取り	アカウントビリティ	プロセス評価、プログラム評価、エンパワメント評価など
②一貫性志向型	管理的な行政技術 —ガバナンス領域内の組織や社会間	一貫性、舵取り	アカウントビリティによるマネジメント	
③管理志向型	管理技術 —組織の内部	（船漕ぎ）	マネジメント	業績測定、評定、政策分析など
④知的貢献志向型	その他の情報提供 —組織と社会	目標設定	専門分野への知的貢献	政策分析、実験手法、経済分析など

（出典：Pierre and Peters 2005：3-5；西尾 1979：40；山谷 1997：67 を参考に筆者作成）

谷 2006: 62-3)。このなかでも山谷は「政策評価」、「ODA 評価」、「独立行政法人評価」の関係を考察している(山谷 2006:257)。日本に「政策評価」の制度が導入された時点ですでに複数の政策評価が存在していたのである。

2019年時点ではさらに評価の種類が増えているので考察対象を限定する必要がある。ODAの行政活動における政策評価の評価システムを考察するにあたって次の4つを取りあげる。すなわち、①各省「政策評価」、②ODA評価、③外交の視点からの評価(以下、外交評価)、④独立行政法人国際協力機構(Japan International Cooperation Agency 以下、JICA)の事業評価(以下、JICA事業評価)である。

4つの制度は異なる評価ポリシーに従っているので、異なる評価システムである。すなわち、「政策評価」は「政策評価に関する標準的ガイドライン」と各省のガイドラインに従っており、組織的な産出情報として「政策評価書」が公開されている。ODA評価は「ODA評価ガイドライン」に従っている。ODA評価の場合には組織的な産出情報としては多くのものがあるが、なかでも「ODA評価年次報告(書)」と「国別評価書」を用いる。外交評価は、ODA評価の制度の一部でもある。そのため、ODA評価と同じ文書が用いられている。しかし、実際は文書内では別の評価として項目を分けて記述されている。JICA事業評価は、「JICA事業評価ハンドブック(Ver.1.1)」と「JICA事業評価ガイドライン(第2版)」に従っており、「国際協力機構事業評価年次報告書」を刊行している。これらの制度は、いずれも組織の責任のもとで評価の形をとっており定期的に刊行されて政策過程に対するフィードバックを意図しているので、すでに検討した評価システムの特徴の要件を満たしている。

4つの評価システムは、評価対象によってすみ分けできているのだろうか。山谷は、政策体系を政策(policy)、プログラム(program)、プロジェクト(project)のヒエラルヒーに分けて、評価対象のレベルをもとに評価と測定を区別している(山谷 2012: 18 図1)。しかし、現在のODAにおける評価システムは、対象の範囲が不明瞭になっている。たとえば、ODA評価には評価対象をセクター別、スキーム別、国別、重点課題別と分けた類型があるが、政策レ

ベル評価とプログラム・レベル評価の双方に対応してしまう。さらに2017年度以降、外務省が実施する無償資金協力の事業完了済みの個別案件のうち、供与限度額10億円以上の案件についてプロジェクト・レベルの評価が行われるようになった(URL1: 6)。「平成30年度外務省政策評価書」では次の7目標で事後評価が記述されている。すなわち、「地域別外交」、「分野別外交」、「広報、文化交流及び報道対策」、「領事政策」、「外交実施体制の整備・強化」、「経済協力」、「分担金・拠出金」である(URL2: 6)。これに加えて、交換公文(Exchange of Notes 以下、E/N)供与限度額10億円以上の無償資金協力プロジェクトおよびE/N供与限度額150億円以上の有償資金協力プロジェクトについて事前評価を行っている。また、5年間未着手の事業および10年経過しても未了の有償資金協力についての事後評価もある。さらに、2016年秋の行政事業レビューによる指摘を受けて、外務省が直接実施する無償資金協力についてもプロジェクト・レベルの第三者評価を実施することになった(URL3: 11)。JICA事業評価もまた、対象の範囲を拡大してプログラム・レベルにまで関与する。このように評価対象のレベルの垣根を超えて重層化しているのである。

だが、評価活動の重層化は無駄な重複ではない。むしろ、他の評価システムの結果が活用されて、新たな評価結果をつくる流れになるメリットもある。そこで、この流れを見ながら評価システムの類型と複雑な関係を明らかにする。

ODA評価はどのような様相だろうか。ODA評価は、管理志向型の特徴もあわせもつアカウントビリティ志向型の評価システムであると考えてみたい。ガイドラインでは「外務省のODA評価においても、ODAの管理改善(ODA政策へのフィードバック)及び国民への説明責任の確保を目的にしている」と記述されているからである(URL3: 2)。この目的の設定は、ODA評価のルーツである経済協力評価の当初の目的が個別プロジェクトの管理支援であったという歴史的理由によるものである(URL3: 13)。

しかも、ODA評価には一貫性志向型の側面も指摘できる。ODAに関係する経済産業省や文部科学省といった他省庁も各々の「政策評

価」あるいは個別評価を行っている。これらの評価結果は外務省によって ODA 評価年次報告書にまとめられている (URL4:27-37)。同様に、JICA 事業評価の結果もまた ODA 評価年次報告書の一部分として利用されている (URL4:38-9)。ODA 評価年次報告書には、外務省の「調整官庁」としての側面を見て取ることができるのである。

長い歴史を持つ ODA 評価に追加されて、いまだ試行錯誤の段階にある新しい試みが外交評価である。これは国際的に見ても特殊な試みであり、理論よりも実務が先行している。制度上は ODA 評価の一部であるものの実態は異なる機能を果たしている。それでは、外交評価は ODA 評価とはどのように異なるのか。外交評価の該当箇所は、ODA 評価の各報告書に新設された「外交の視点からの評価 (Diplomatic Viewpoints)」である。その代わりに、旧来の ODA 評価は、「開発の視点からの評価 (Development Viewpoints)」として区別された。このように、組織活動を規定する文書や評価結果の項目が分けられている特殊性から、制度上は ODA 評価の一部とされていても別の評価システムとして扱うのが妥当である。

外交評価は、ほかのすべての評価システムの結果を利用している。外交評価は ODA 評価の結果の活用するようにガイドラインは定めている (URL1:22)。先に考察したように「政策評価」や JICA 事業評価の結果を ODA 評価は活用しているので、外交評価はこれらの評価結果も間接的に利用していることになる。

具体例として、「平成 30 年度インドネシア国別評価書」の外交の視点からの評価項目から視点を抽出する (URL5:URL6)。開発の視点からの評価が、政策の妥当性、結果の有効性、プロセスの適切性に焦点を当てているのに対し、「外交的な重要性 (Diplomatic Importance)」と「外交的な波及効果 (Diplomatic Impact)」に着目して記述的に評価を行っているとわかる。

ODA 評価と異なる特徴は、日本の国益という視点を強調している点である。たとえば、被援助国において日本に対する信頼が醸成されているのか、その開発援助は外交カードになるかなど国内向けのアカウンタビリティ確保を目的としているアカウンタビリティ志向型の評価システムに類型化できる。ただし、「外

交 (Diplomatic)」とはいっても外交官が行う対外交渉や外交儀礼といったプロフェッショナルな外交を評価対象にするわけではなく、おもに ODA が外交に貢献しているかどうか国別開発協力方針との整合性を通じて確認しているにとどまる。

むしろ、外交をとらえようとしているのは外務省の「政策評価」かもしれない。外務省の政策評価書の各項目には、ODA 評価や外交評価と類似した指摘が見つかる。たとえば、「経済協力」の項は ODA 評価に重なり、「地域別外交」の項は外交評価と重なる。先述した「インドネシア国別評価書」において「外交的な重要性」としてあげられた経済連携協定に関する議論は政策評価書にも同様に記されている。

他の評価システムの結果を利用しているのは、「政策評価」も例外ではない。「測定指標 7 主要個別事業の事後評価結果」の項には、JICA の事業評価年次報告書の評価結果を利用していることが明記されている (URL2)。「政策評価」の未着手・未了案件の事後評価でも「社会ニーズの現状」と「事業遅延に関する経緯・現状」の 2 点の確認を通じた評価のなかで、おもに JICA の事業事前評価と E/N の情報が利用されている (URL2:615)。事業事前分析表を用いた目標管理型へと移行した「政策評価」は、管理的な側面を強調するために他の評価結果の活用をしており、管理志向型の側面を強めている評価システムであるといえる。

他の評価システムの随所に基礎的な情報を提供しているのは、JICA 事業評価であった。その目的は「JICA 事業評価ガイドライン (第 2 版)」によれば、①事業の実施監理を改善するための情報提供と②「日本国民および相手国を含むその他ステークホルダーへの説明責任 (アカウンタビリティ) の確保」 (URL7:2) であると記述されており、管理志向型の側面が強い。そのうえで、ODA 評価の部分としてのアカウンタビリティ志向型の側面もあるが、外交評価とは異なってより広い対象者を想定しているのが特徴的である。

以上の考察から次の 2 点がわかる。第 1 に、異なる評価システムが互いに重層化しながら複雑な関係にあることで全体の評価システムを作り上げていた。第 2 に、この複雑な実態には、管理技術的な評価システムの結果が行政技術的

な評価システムへと活用されていく一連の流れがみられた。すなわち、JICA 事業評価は「政策評価」と ODA 評価と外交評価に対して、「政策評価」は ODA 評価に対して、ODA 評価は外交評価に対して産出情報をフィードバックしていた。

このような産出情報のフィードバックによって、ガバナンス機能確保のための中心的な要素であるアカウンタビリティ追及のメカニズムは、全体として一定程度確保されていると考えられる。そして、外交評価などの新しい評価システムの導入はアカウンタビリティの変化に対する適応なのであり、ガバナンスの確保に対する外務省の努力の成果であるといえる。

5. おわりに ——評価システムのガバナンス

本稿には次のような限界がある。第1に、評価システムの理論的類型化作業にはいまだ他の視点による改善の余地がある。第2に、ガバナンスの定義が比較的限定的であった。したがって、他のガバナンスを考察する場合には別の視点が必要である。第3に、組織文化や非公式なルールの影響も評価システムにあると強調したものの、本稿では扱いきれていない。

以上のような限界はあったものの、本稿の目的は次のように達成できた。第1に、評価システムという新しい枠組みによって評価の複雑な実態を把握する方法を提案した。この枠組みを整理した結果、行政学と政策学と評価学の視点を実務の関心のもとに結びつけるものであるとわかった。つまり、日本で培われてきた学問の延長線上にある概念であった。第2に、評価システムの理論的類型をガバナンス機能と関係させて導出した。すなわち、「アカウンタビリティ志向型」、「一貫性志向型」、「管理志向型」、「知的貢献志向型」の4つの理論的類型を提案した。第3に、日本の評価システムの一例として、ODAの行政活動における政策評価の複雑な実態を分析した。その結果、各々の評価システムが生み出す情報の相互利用を通じて重層で複雑な評価システムが描出された。

なお、想定しなかった発見もある。ODAにおける評価の複雑な実態はむしろガバナンス機

能の確保に貢献していた。すなわち、管理志向型の評価システムの結果がアカウンタビリティ志向型の評価システムによって利用されることで、全体としてのガバナンス機能の確保に貢献していた。ようするに、評価の氾濫それ自体が問題であるとはいえない。

ただし、この複雑な実態には課題も残る。すなわち、実務者や評価の専門家には理解可能であるが、外部の市民や住民にとっては難解であり混乱するという複雑さの課題は解決されていない。なぜなら、増える評価と飛び交う評価の産出情報を統制する舵取りが存在しないからである。それでは本来の意味でのアカウンタビリティが確保されておらず、ガバナンスの有効性も損ねてしまう。したがって、評価の氾濫をガバナンスの確保に活かすためには、複雑化する評価システムそれ自体のガバナンスが必要になるだろう。

評価の氾濫には、評価システムのガバナンスが有効な処方箋になりうる。この処方箋を考えるうえで生じる3つの課題が今後の展望である。第1に、行政過程における評価システムの影響である。すなわち、本稿では扱いきれなかった独立行政法人評価や行政事業レビューなどはどのような影響をガバナンスに及ぼすのだろうか。第2に、評価システムの通時比較である。すなわち、本稿で説明した現状に至るまでの評価システムの変容過程はどのようなものであり、どうやって変容を舵取ればよいのだろうか。第3に、評価システムの共時比較である。すなわち、他国や他政府、他の分野の評価システムにはどのような特徴があるのだろうか。本稿で論じた枠組みによって、このような比較評価の考察が可能となり、評価の氾濫に対応するための道が開かれる。

参考文献

- 日本語文献
秋吉貴雄・伊藤修一郎・北山俊哉（編）（2015）『公共政策学の基礎 新版』有斐閣。
今村都南雄（1997）『行政学の基礎理論』三嶺書房。
城山英明（2007）『国際援助行政』東京大学出版会。
西尾勝（1979）「管理技術の発展と行政」『季刊行政管理研究』6、40-45。
西尾勝（1990）『行政学の基礎概念』東京大学出版会。

橋本圭多 (2017) 『公共部門における評価と統制』 見洋書房。
 森田朗 (1988) 『許認可行政と官僚制』 岩波書店。
 山谷清志 (1997) 『政策評価の理論とその展開—政府のアカウントビリティ』 見洋書房。
 山谷清志 (2006) 『政策評価の実践とその課題—アカウントビリティのジレンマ』 萌書房。
 山谷清志 (2008) 「ODA 評価と政策評価—日本の現状分析」 湊直信・藤田伸子 (編) 『開発援助の評価とその課題』 51-73、財団法人国際開発高等教育機構。
 山谷清志 (2012) 『政策評価』 ミネルヴァ書房。

英語文献

Considine, M. and Afzal K. A. (2012) Accountability in an Age of Markets and Networks. In Pierre, J. and Peters, B. G. (eds.) *The SAGE Handbook of Public Administration*, concise second edition, 615-633, Sage Publications.
 Furubo, J. E. Rist, R. C. and Sandahl, R. (eds.) (2002) *International Atlas of Evaluation*, Transaction Publishers.
 Hanberger, A. (2011) The Real Functions of Evaluation and Response Systems. *Evaluation*, Sage Publications, 17(4), 327-349.
 Højlund, S. (2014) Evaluation Use in Evaluation Systems: The Case of the European Commission. *Evaluation*, Sage Publications, 20(4), 428-446.
 Kjær, A. M. (2004) *Governance*, Polity Press.
 Leeuw, F. L. and Furubo, J. E. (2008) Evaluation Systems: What Are They and Why Study Them?. *Evaluation*, Sage Publications, 14(2), 157-169
 Liverani, A. and Lundgren, H. (2007) Evaluation Systems in Development Aid Agencies: An Analysis of DAC Peer Reviews 1996-2004. *Evaluation*, Sage Publications, 13(2), 241-256.
 Martinaitis, Ž., Christenko, A. and Kraučūnienė, L. (2018) Evaluation Systems: How Do They Frame, Generate and Use Evidence?. *Evaluation*, Sage Publications, 21(1), 1-16.
 Palfrey, C., Thomas, P. and Phillips, C. (2012) *Evaluation for the Real World: The Impact of Evidence in Policy Making*, The Policy Press.
 Pierre, J. (ed.) (2000) *Debating Governance: Authority, Steering, and Democracy*, Oxford University Press.
 Pierre, J. and Peters, B. G. (2000) *Governance, Politics and the State*, Macmillan Press.
 Pierre, J. and Peters, B. G. (2005) *Governing Complex Societies: Trajectories and Scenarios*, Palgrave.
 Raimondo, E. (2018) The Power and Dysfunctions of Evaluation Systems in International Organizations, *Evaluation*, Sage Publications, 24(1), 26-41.
 Rist, R. C. and Stame, N. (2006) *From Studies to Stream: Managing Evaluative System*, Transaction Publishers.
 Vedung, E. (2010) Four Waves of Evaluation Diffusion, *Evaluation*, Sage Publications, 16(3), 263-277.

URL

1. 外務省大臣官房 ODA 評価室 (2018) 「ODA 評価ガイドライン 第 11 版」 (2019 年 11 月 25 日取得、<https://www.mofa.go.jp/mofaj/gaiko/oda/files/000369136.pdf>)
2. 外務省 (2019) 「平成 30 年度外務省政策評価書」 (2019 年 11 月 25 日取得、<https://www.mofa.go.jp/mofaj/files/000394378.pdf>)
3. 外務省大臣官房 ODA 評価室 (2019) 「ODA 評価ガイドライン 第 12 版」 (2019 年 11 月 25 日取得、<https://www.mofa.go.jp/mofaj/gaiko/oda/files/000486015.pdf>)
4. 外務省 (2018) 「ODA 評価年次報告書 2017」 (2019 年 11 月 25 日取得、<https://www.mofa.go.jp/mofaj/gaiko/oda/files/000325965.pdf>)
5. 国際航業株式会社 (2019) 「インドネシア国別評価 (第三者評価) 報告書」 (2019 年 11 月 25 日取得、<https://www.mofa.go.jp/mofaj/gaiko/oda/files/000496682.pdf>)
6. Kokusai Kogyo Co., Ltd. (2019) Evaluation of Japan's ODA to the Republic of Indonesia, (Retrieved on November 25, 2019, <https://www.mofa.go.jp/policy/oda/evaluation/FY2018/pdfs/indonesia.pdf>)
7. 独立行政法人国際協力機構評価部 (2014) 「JICA 事業評価ガイドライン (第 2 版)」 (2019 年 11 月 25 日取得、https://www.jica.go.jp/activities/evaluation/guideline/ku57pq00001pln38-att/guideline_ver.02.pdf)