

第2章 JICAの評価と事業の改善に向けて（提言）

外部有識者事業評価委員長 同2次評価作業部会部会長 牟田博光

【2次評価の新しい試み】

2003年度に引き続き、2回目の2次評価結果が公表された。昨年度、外部有識者にはそれぞれ評価傾向があることを明らかにし、偏りのない評価結果を得るためには複数の評価者に評価させて比較することが重要だということを実示した。今年度は前回と比較して、経年的な変化を調べる、外部有識者だけでなくJICA職員も参加して2次評価を試みる、などの新しい試みも行った。

もし評価結果が信用できないということであれば、そのフィードバック効果は限られる。すべての人が納得できる評価結果はないかもしれないが、できるだけ多くの人が納得できる評価結果を得る仕組みは作ることができる。2次評価はそのような仕組みとして考えられた。

JICA職員と外部有識者の比較は面白いと同時に有益である。どういった視点が違うのか、なぜそうなのかに関して、今年度の結果からだけではすべてが明らかになったわけではないが、このような試みを続ける必要がある。JICAには有能な人材が沢山いる。評価においても、JICA内部の人的資源を有効に活用することがまず必要である。

JICA職員が外部有識者と一緒に2次評価作業を行うことで、外部者の目で評価を行うことを理解してもらい、また有識者もJICA職員と一緒に評価作業をすることで、独り善がりに陥らない評価の目を養うことができよう。評価において、外部有識者とJICA職員とは立場や経験が違うだけで、どちらが優れているということはない。お互いが相手の視点を理解することが評価者個々人の評価の目を肥やすことにつながると同時に、評価にふくらみをもたらす。

【評価者の評価】

評価者の評価傾向がわかるということは、評価者が評価されるということでもある。有識者に限らず、評価者に一定の評価傾向があること自体は悪いことではない。評価者の専門分野、経験のみならず価値観も評価傾向に影響を与えるのはある意味では当然である。しかし、同じ評価者が同じ対象を評価すればいつも同じ評価結果が

得られることが望ましい。評価するたびに評価結果が変わるようであれば、結果が信用できないことになる。現実には同じ案件を何度も評価することはないので、このようなことはわからないが、ここで用いたような方法で一般的な評価傾向を算出すると同時に、評価傾向そのものを、真の評価傾向と誤差とに分け、誤差が大きく、評価傾向が大きくふらつくのであれば、評価結果が信用できないことになる。このような試みを続けることによって、評価者としての適切性を磨いていくことも可能であろう。

【評価の現場主義】

新ODA大綱や新ODA中期政策では現場主義が強調されているが、現場主義とは案件の形成や実施だけではなく、評価活動も含めて考えなければならない。しかし、評価の多くが在外事務所主導で行われるとしたら、その出来映えをどこかでチェックして、質を維持する仕組みがなければならない。評価が重要になればなるほど、その品質保証が必要である。この2次評価手法を用いれば、評価の質を本部でチェックする仕組みを構築することができよう。もし評価の質に問題があるとすれば、その原因は案件なのか評価者なのか、あるいは事務所の体制なのかの判断も可能であろう。

【アウトカムの重視】

案件が成功したかどうかを評価する判断基準として重要なのは、アウトカムの発現であるが、2次評価分析の結果からは外部有識者とJICA職員との判断の差が大きい評価視点である。PDMに即して言えば、高次の目標が実現したかどうかの視点である。アウトプットがなければアウトカムは発現しないが、アウトプットがあってもアウトカムが自動的に発現するわけではない。アウトカムの発現にはある程度の時間がかかるという事情もあるが、それよりもアウトプットからアウトカムへ至る道筋が相対的に細かいことの方が問題である。アウトカムに寄与するアウトプットは通常1つではないし、寄与の程度も

様々である。元々寄与の程度が小さいのであれば、時間が経っても目に見えるアウトカムの変化は得られない。それに、時間が経てば経つほど案件以外のアウトプットの影響も強くなり、アウトカムに変化があったとしてもそれが案件のアウトプットの影響であることを証明することは困難となる。時間が経てばアウトカムが発現するという事は理屈としてあっても、時間がたてばたつほど、実際にその寄与を計測することは期待できない。

案件は通常3～5年は継続している。案件終了時には案件初期の活動がいくらかのアウトカムの芽を生み出していることは予期できることであり、それを上手に見いだすことが大事である。アウトプットからフルスケールのアウトカムまでには距離があるにしても、その途中の様々な中間指標を拾い上げることによって、アウトカム発現へのより確かな見通しを描くことはできる。評価報告書の中に、アウトカムについて「まだ発現していない」と書くのではなく、アウトカムに向けてどのような芽が出ているかについて詳述するように心がける必要がある。

【プログラムアプローチの評価】

これまで、「プロジェクト活動→アウトプット→アウトカム→インパクト」という一連の筋のおった因果関係の見通しをもって案件を実施することが重要とされてきた。またそのような順序で評価するのが良いとされてきた。それは今でも間違っていないが、原因-結果の見方を逆にして、インパクトを得るためにはどのようなアウトカムが必要か、アウトカムを得るためにはどのようなアウトプットが必要か、と逆にたどっていけば、道筋は1本では済まなくなる。アウトカム重視の立場から、必要とされるインパクトやアウトカムは何かをまず定めて、それに強く貢献することが期待できるように案件を形成するような、多くの要因を構造的に見通したプログラムの視点が案件形成に求められている。このように、プログラムの視点から原因-結果の複合的連鎖によってインパクトを確実に達成するようプロジェクトを構築するのであれば、評価を行う際にもプログラムの視点の評価が必要となってくる。

また、これまでPDMなどで外部条件としてひとくくりにされてきた要因の中に、本来ならばプロジェクトの一部として取り組むべき要因が含まれている場合も多い。そうであれば、当初案件の作り方にかかわらず、評価の際にはプログラムの視点から、多方面に目配せのきいた評価を心がけることが重要である。

【事業改善のツール】

評価にはコストがかかるが、それだけのコストをかけて何がどれだけ改善されたかも評価されなければならない。評価の目的は改善にある。今回の分析で評価報告書の質が向上したことが示されたが、評価ガイドラインの作成・徹底をはじめとする近年の取り組みによるところが大きいと思われる。また、JICAでは2001年以降、事前から事後までの一貫した評価体制を導入し、2004年2月にはガイドラインを改定しているが、これらの取り組みが今後どのように評価の質の変化に反映されてくるかが注目される。

本報告書で用いたような評価視点に留意することをおして、そもそもどのように案件を形成すれば高い成果が得られやすいかの指針も出てくる。何が評価されるかという視点を理解して、その達成に注意をはらって計画すれば、自ずと案件形成の質は向上する。この作業部会で開発した2次評価票は案件形成を行う際のチェックリストとしても使える。どのような視点で最終的に評価されるかわかっているのだから、それらに注意して案件形成すれば成果が得られやすいからである。同様に、案件実施上の留意点を分かりやすく示すことによって、案件の成果そのものを高めていくこと、またそれを実証していくことも望まれる。

眼光紙背に徹すれば自ずと改善点が見えてくる、という報告書では実用的ではない。評価報告書を基にして、案件形成や事業実施のための手引き、チェックシートなどを作り、誰でもが評価結果を利用しやすい形に情報を加工、発信することが今後ますます重要である。

【最後に】

評価は事業の質保証のための仕組みである。事業の質の向上のためには、職員一人一人の心がけが大事であることはもちろんであるとしても、質を担保する仕組みを構築し、心がけを支えることもまた重要である。本報告書のような評価活動はまだ2年目であるが、上記のようなアンビシャスな評価を目指して、これからも努力を継続して頂きたい。

2004年度2次評価結果を受けて

—2次評価結果の有用性とJICAにおける活用

JICA評価検討委員会委員長 理事 小島誠二

今般、58件に及ぶ終了時評価を対象に、より良い手法を検討のうえ、知見を生かして信頼性・納得性の高い評価を実施していただいた、外部有識者事業評価委員会2次評価作業部会の評価者各位のご尽力に心より感謝いたします。

JICAは2002年度に外部有識者事業評価委員会を設置し、同委員会による2次評価を導入しました。2次評価は、JICAが評価の透明性と客観性を高め、より適切に説明責任を果たしていくうえで重要なものと考えています。また、評価の質の向上に向けた示唆を得るうえでも大変に有益なものです。この認識のもとで、昨年度に実施していただいた2次評価については、その結果を『事業評価年次報告書2003』への掲載を通じて公表するとともに、貴重なご指摘・ご助言を活用し、2004年2月に「事業評価ガイドライン」の大幅な改定を行いました。そして現在、改定ガイドラインにもとづく評価の徹底と評価の質の改善に取り組んでいるところです。

本年度の2次評価では、JICAの評価の質や事業に関し、より掘り下げた分析を通じて多くの貴重なご示唆をいただきました。評価や事業の質が一定のレベルを満たしているとの評価結果、とりわけ評価の質が経年的に向上しているとの評価結果は、評価や事業の改善に取り組んできたJICAにとっては嬉しいことです。しかしながら、その一方で、新たな事業へのフィードバックのための「提言・教訓」の質の低さをはじめ、「効率性」や「インパクト」の視点の弱さ、途上国側の評価への参加度や受益者からの情報収集の不足など、まだまだ取り組むべき課題が少なくないことも示されています。JICAでは、いただいたご指摘をふまえ、引き続き、評価と事業の改善に取り組んでいきます。

その一環として、ご提言にありますように、2次評価を今後とも継続・拡充するとともに、今回、提案いただきました2次評価チェックシートをJICAにおける評価の運営管理に活用し、1次評価の質の改善を図っていくつもりです。

また、現在、JICAではより効果的な事業の実施に向けて、「現場主義」を進めており、在外事務所が事業の発掘・形成、計画、実施、モニタリング、評価に中心的な役割を果たしつつあります。そうした中、在外事務所の評価実施体制の強化や評価の質の向上は「現場主義」

の推進に向けての重要な課題の1つとなっています。2次評価は、在外事務所による評価の質の管理、そして評価を通じた事業改善に有益であり、体系的な活用を検討していきたいと考えます。

さらに、案件形成にあたっての「評価の視点」の意識は、結果重視の事業運営に向けて欠かせないものであり、そうした観点からの2次評価チェックシートの活用についても、あわせて検討していくつもりです。

なお、結果やインパクトの確実な達成に向けた、案件形成や評価に際してのプログラムの視点の重要性に鑑み、現在、JICAでもプログラム・アプローチを強化しているところです。具体的には、事業の計画・実施面では、これまでは概念のみであったプログラムを実際の事業運営の単位として、プログラム・レベルで予算を配分したうえで運営管理する制度を2004年度から導入しました。また、評価においても、これまでの国別事業評価やプログラム・レベルの評価の経験をふまえ、事業のプログラム化に対応し、プログラムを対象としてより質の高い評価を行っていくために、現在、評価手法の改善および評価を活用したプログラムの改善に取り組んできました。JICAでは、外部有識者事業評価委員会のご助言もいただきつつ、これらの作業に引き続き取り組んでいきたいと思えます。

最後に、JICAでは、評価の質の改善のためには「良い評価」の例を具体的に示すことが有効という、昨年度の2次評価でのご助言を受けて、2004年度より「JICAグッド・プラクティス評価賞」を導入いたしました。そしてその中で、他の模範となるような、評価結果のフィードバックの取り組みや質の良い評価をグッド・プラクティスとして選び、広く共有のうえ、組織レベルでの学習を行っていくことにしています。第1回の「JICAグッド・プラクティス評価賞」では、質の良い評価を対象とする「優秀評価賞」として、2次評価において「良い評価」とされた4件の終了時評価報告書を選びました。JICAは今後とも、2次評価結果を活用して質の良い評価を客観性の高い形で選り出し、「JICAグッド・プラクティス評価賞」を通じてその内容を広く共有することで、模範となる評価の蓄積を図り、評価の質の向上につなげていきたいと考えています。

テクニカル・ノート

*1 評価得点の推定

各人が評価していない案件の評価項目および小項目の評点推定に際しては、案件と評価者のダミー変数を説明変数、個人の評価得点を従属変数とした以下の回帰式で求められる重回帰分析を行った。二次評価チェックシートのすべての評価項目および小項目についても同様の処理を行った。

$$\hat{Y}_{ijk} = \alpha + \beta i - 1D_{p,i-1} + \gamma k - 1D_{e,k-1}$$

\hat{Y}_{ijk} = チェックシートの評価項目および小項目の得点の推定値 (9の評価項目と43の小項目)

($i=1, \dots, 48$, $j=1, \dots, 52$, $k=1, \dots, 14$) (i は案件数、 j は評価項目・小項目数、 k は評価者数)

$D_{p,i-1}$ = 2002と2003年度の計48案件のダミー変数 ($i=1, \dots, 48$) ($D_p=1$, または $D_p=0$)

$D_{e,k-1}$ = 評価者のダミー変数 ($k=1, \dots, 14$) ($D_e=1$, または $D_e=0$)

$\gamma k - 1$ は各評価者の評価傾向を示す「評価傾向係数」である。平均の定義により平均値と各データの値の差の合計は0となることから、評価傾向係数の合計を0になるように調整することにより、評価項目および小項目の推定値が、全員が評価した際に得られると考えられる平均値となる。すなわち、

$$\sum_{k=1}^{14} \gamma k - 1 = 0$$

また、2001年度については2人のみが評価しているが、2002年度と2003年度の評価傾向係数より、残り12人が評価した場合の値を推定した。

なお、ダミー変数は質的な現象を回帰分析に反映する手段である。ダミー変数はカテゴリー数から1を引いた個数だけ必要となり、その値は0, 1の組合せとなる。また、回帰分析とは、ある要因Xが他の要因Yにいかほどのインパクトを与えるか、与えたとしたらどの程度かを分析する手法である。重回帰分析とは要因Xが2つ以上ある場合の回帰分析を呼ぶ。

*2 評価項目間の相関関係

評価項目間の相関関係は、相関係数を用いて分析した。相関係数とは、量的変数間の関係を探る場合に使用する統計量である。相関係数は-1から1の間の値をとり、|1|に近づくほど相関が強くなる。また、-1に近づくほど変数間は負の相関関係、1に近づくほど正の相関関係があることになる。

今回の2次評価の評価項目間の相関関係は、すべて相関係数1%水準で有意(両側)となっている。特に「情報収集」と「分析/評価」の

全項目(現状把握、分析、評価)、「分析/評価」の3つの項目間、「現状把握」「分析」と「提言・教訓」、そして「分析/評価」(全項目)および「提言・教訓」と「総合評価」の間は、相関係数0.7以上で強い相関関係が示されている。また、「総合評価」は「評価フレームワーク」「情報収集」「報告書」とも相関係数0.6以上で強い関係を示している。

*3 年度間の質の差の検証

年度間の質の差の検証では、分散分析および多重比較を用いた。分散分析は、3つ以上の母平均の間の差の検定を行う場合に用いる分析手法、多重比較は、分散分析においてどのグループ間に差があるかを検定する手法である。分散分析では、有意差を出しやすいとされるLSD法を使用した。そして年度間の母集団の平均点に差がない確立が5%以下(有意確率が0.05以下)のものについては有意である(95%の確率で平均点に差がある)と判断した。

*4 外部評価者と内部評価者の評価結果の差の検証

外部評価者と内部評価者の評点の差の検証は、母平均の差の検定を用いた。差の検定は、母集団の平均がある値と等しいかどうか、また、2つの母集団の平均が等しいかどうかを検定する方法である。

*5 馴染み度の評点への影響の分析

小項目、年度、馴染み度を独立変数、評価項目を従属変数として、評価者に回帰分析を行った。回帰分析とは、ある要因Xが他の要因Yにどのぐらいのインパクトを与えるか、与えたとしたらどの程度かを分析する方法である。

*6 評価傾向の違いの分類

評価傾向の違いの分類は、因子分析を用い、得られた因子得点について評価者ごとに平均点を出したのち、クラスター分析を行い、回答者をグループ分けした。因子分析は、多くの変数がある場合にそれらの変数の共通性を抽出し、その共通因子でデータを説明する方法であり、因子得点は各オブザベーションごとに計算される因子の得点である。また、クラスター分析は個々のサンプルや変数をいくつかの似たもの同士のグループに分類する方法であり、図4-8はクラスター分析のWard法を用いている。

*7 小項目の重み付けに関する統計分析案

評価項目を従属変数、小項目を独立変数として、標準回帰係数を用いて重回帰分析を行い、重回帰分析の係数に、上記6の因子分析の結果を加味した。標準回帰係数は、すべての係数を標準化したのちに回帰分析を行い得られる係数である。

*8 評価項目の精選

小項目間の相関が特に高く、かつ因子分析の結果、同一因子について因子負荷量が高い(絶対値で0.4以上)小項目を選んだ。因子負荷量は、因子の各変量への影響度である。