

教育援助プロジェクトの事後評価 理論的裏付けと実践的課題⁽¹⁾

長尾 眞文

(広島大学教育開発国際協力研究センター)

1. はじめに

教育援助は、開発途上国に対する人づくり援助の中核領域としてODAが本格化した1960年代当初から常に優先的に扱われてきた。しかし、その優先的性格とは裏腹に(あるいはそれゆえに)長い間教育援助の評価が重要検討課題として取り上げられることはなかった。例え評価関心があったとしても、援助実施プロセスのモニタリングとマクロ的・平均的統計による結果のチェック程度(例えば、就学率や識字率)であった。教育援助の評価、とりわけ事後評価により高い関心が集まるようになったのは、1990年代後半からで、その背景には対途上国援助の効果に関する一般的関心の高まりがあった。ドナー諸国政府が1990年代にほぼ例外なく採用した成果管理主義の影響で、援助の成果重視志向が強まり、評価に対する関心が顕著に拡大した。具体的には、援助評価システムの整備・拡充が進むとともに、援助評価を事前・中間・事後とサイクル的に捉えるアプローチ、ドナー間で合同評価する試み、途上国の評価への参加など、多様な評価の取組みが試みられた。DACが1996年に発表した「新援助戦略」で、ドナー諸国が開発目標設定型の援助を謳ったことも、事後評価重視の流れを後押しした。

一方、教育援助の分野においても、事後評価関心の覚醒につながる展開が見られた。第一に、1990年の「万人のための教育(EFA)」世界会議(ジョムティエン会議)を契機として、世界的に未就学児童の減少や初中等教育

における男女間格差解消の目標が共有されるに従い、その目標の達成度のモニタリング・評価が国際社会の重要関心事となった。また、多数の二国間・多国間援助機関が合同で基礎教育分野におけるドナー支援の効果を検討する試みが企てられたりもした(Netherlands Ministry of Foreign Affairs 2003)。第二に、援助全般に亘って開発効果の持続可能性が問われるようになり、それに関連して途上国のキャパシティ・ビルディングの取組みが重視されるようになったことで、派生的に教育・訓練・研修活動の評価に対する関心が高まってきた(Chapman & Nagao 2006)。第三に、日本の途上国協力において近年教育協力の比重が急速に高まってきたことが挙げられる。それ自体は世界規模でのEFA支援の流れに押された動きであるが、日本の援助総額が減少傾向にある中で教育援助の増加であり、教育援助内容の多様化や教育NGOの関与の拡大等の事情も相まって、援助機関には通常以上の評価圧力が掛かっているように見受けられる。

そのような状況認識のもとで、本稿では教育援助評価の理論的裏付けについて簡単にレビューするとともに、その実践的課題について整理を行うこととしたい。前者については、評価の理論的枠組みの構築と実践への橋渡しの双方における有用性で各国の援助機関や国際機関に採用されているロジック・モデルを主として扱うことにする。後者については比較的完成度の高い事後評価システムを確立している世銀の慣行を例にとりて検討することにする。

2. 評価理論の基礎

Shadish, Cook & Leviton (1991)によると、評価の理論とは評価方法の選択の理論のことで、個別の事業の評価をいかに行うかの方法に関する系統的理解と表現できる。具体的には、評価結果の活用の理論(誰のために、何を目的に評価を行うか)、事業計画の理論(事業の想定している因果関係をどう捉えるか)、判断に関わる価値の理論(判断基準を何に求めるか)といった構成要素を組み合わせることで体系化される。中でも個別の評価作業の方向付けに最も大きく影響するのは、評価の目的の設定が評価理論の基礎にあると考えてよい。本節では、この評価目的の理論的扱いについて取り上げる。事業計画の理論は、評価手法に関係し、理論と実践の橋渡し役を果たすもので、次節で取り上げることにする。はそもそも評価対象とする事業や評価行為を支える価値基準に関係する本源的要素で、本稿では最終節で触れるに止める。

Patton (1997)は、評価を行う目的(特に、誰が評価結果を活用するか)によって、評価作業を下記の通り「判断指向型」、「改善指向型」、「知識創出型」の3つのタイプに分ける。

(1) 判断志向型評価

評価対象となる事業について、主として説明責任の視点から事業の価値や扱い(例えば、完了、中止、継続、改善)について総括的判断を行うために実施する評価。総括評価がその典型で、行政監査、会計監査、費用便益分析、品質管理も広義にはこのタイプの評価と考えられる。事後評価で通常想定されるのは、このタイプの評価である。この種の評価で最も重要な点は、判断基準の特定で、例えば学校建設プロジェクトのインパクトを何を尺度として評価するかである。

(2) 改善志向型評価

事業の継続を前提として、主として発展的視点から改善案の提示を行うために実施する評価。形成評価がその典型で、継続的学習、SWOT分析、品質改善等がこの範疇に入る。このタイプの評価では、改善提案の妥当性が重要であることはもちろんであるが、改善の実現可能性は事業管理者をはじめ事業に関与する関係者の当事者意識に掛かる部分が大いなので、評価の実施プロセスをどう組むかが重要な意味を持つ。

(3) 知識創出型評価

共通性のある複数の事業を対象に、学術的あるいは管理的視点から一般的な知識・理解を導くことあるいは理論を構築することを目的として行う評価。横断的・総合的分析が特徴。援助機関が何年間か援助を行った分野で、支援経験からの教訓の抽出を行うことが典型例として挙げられる。そのような評価の結果は、将来的な援助プログラムの企画・運用や援助政策の見直し等に活用される。

援助機関が実施するプロジェクトの事後評価では、上記の3つの目的別評価タイプが多様な形で使われる。例えば、JICAにしてもJBICにしても、個々のプロジェクトの完了時および一定期間後にそのプロジェクトの達成した成果あるいはインパクトについて説明責任の遂行の視点から判断指向型評価を行うことが慣行化している。教育開発援助の例を挙げるなら、JBICの円借款による小学校建設のインパクト評価がそれである。JICAの途上国理数科教育支援がそうであるように、教育分野における技術協力プロジェクトはフェーズを重ねることが多いが、フェーズ間で改善指向型評価が実施されるのが常である。また、近年、国際協力機構(JICA)は理数科教育や高等教育等の分野で過去に実施したプロジェクトの評価を横並びにして再検討することにより総合分析評価を行っている。(国際協力機構2004a、2006)これは過去の

教訓を将来の事業実施に生かすための知識創出型評価の典型的な例である。

3. ロジック・モデルの活用による理論と実践の橋渡し

教育分野であれど、開発分野であれ事業の事後評価の主要な関心は、事業計画に盛り込んだ取組・介入の実施により想定通りに成果が生まれ、当初立てた目標を達成できたかを検証し、もし達成できなかった場合には、なぜできなかったかを説明することである。事業、プロジェクトは一定期間内の取組であるから、その効果を見るには事前と事後の状況をなんらかの形で比較しなければならない。事後評価といえども、正確に評価しようとするとは事前（つまり事業が始まる前）に事業計画の中に組み込まなければならない。

正確に評価するとは、例えば、事業で得られた成果が想定した因果関係によるものであるかどうかを判断するのに、「取組を行わなかった場合と比較してどうであるか」、「プロジェクトの外部要因の影響はなかったか」、さらには「データの選択が適切であったか」といった問い掛けを行うことを意味する。社会科学的研究の方法論として最も科学的に厳密であるとされる「実験デザイン」アプローチに依拠すると、事業を実施しなかった場合と比較してどうであるかを見るには、そこから受益するグループ（受益グループ）の他に受益しない同質的なグループ（比較グループ）を設けて、取組成果に関する事前と事後のデータを両者について求め、受益グループに起こった変化から比較グループに起こった変化を引いて、プロジェクトの「純粋な」効果を求める（Campbell & Russo 1999）。両グループを出発点で厳密に同質的にしようすると、すでに存在している2グループには何らかの偏りが存在するかもしれず、それぞれのグループに含めるメンバーを同じ集団の中からランダムに同数選ぶ等々、データ処理

についても事前の取り組みが必要となる。正確、厳密な評価を期するには、事業計画の理論的枠組がしっかりしている上に、評価計画が事業開始前から用意されていなければならない。

しかし、実際の事業計画の作成は、時間的制約、不十分な事前調査等により、必ずしも適切に組まれているとは限らない。とりあえず予算確保のために作成した企画素案がそのまま事業計画となったり、課題の重要性のゆえにとにかく事業を起こして「走りながら」計画を立てるといった形でスタートすることも少なくない。事業計画作成プロセスに恣意的な政治的介入があつて本来適切に設計された事業を歪めてしまうこともあろう。事業実施体制が突然の人事異動で変更されたり、事業の環境の条件に急激な変化があつたりすることも珍しくない。実際の開発援助の評価では、ランダム・サンプルの実験デザイン・アプローチに代表されるような理論的配慮が貫徹されるケースはごくごく限られる。

開発現場でより一般的な状況は、事業による取組・介入と期待される成果の間の因果関係の想定が不明確なことである。例えば学校教育では、当然毎日の授業の積み重ねが一定の教育効果を挙げることを想定して教育を行っているが、教員の中には10年1日のごとく同じ授業を繰り返し、期末テストで生徒の成績をつけるだけで、どれだけ教育効果が挙げたかを真剣に考えない、そもそも毎日の教育行為を計画の対象として捉えない者もいるであろう。また、学校建設プロジェクトで、学校を必要としている地域と実際に建てられる地域が異なったり、学校建設の目処は立ったが教員手配が追いつかないといった不適切・不十分な計画事例もあろう。さらには、現職教員研修プロジェクトで、開催した研修会の回数を成果としたり、教員の教科指導力の向上を成果と特定するが成果の測定は行わない、というように計画は立てるが、系統的な評価に耐えない場合も考えられる。

しかし、開発に携わるものであれば誰でも、計画段階で取組・介入と成果の間の因果の想定を明確にすることが望ましいことには異議をはさまないであろう。それができると事業を支える理論的枠組が備わることになるからで、この枠組は事業の実施を下支えするだけでなく、そのまま事業の進捗状況の把握や期末の評価にも活用できる。例えば、事業実施段階では取組経過のチェックや中間的成果のモニタリングに(プロセス評価あるいは実施評価)そして事業終了時や事後的には通常のDAC評価5基準(妥当性、有効性、効率性、インパクト、持続可能性)を使った達成度の評価(成果評価)に活用すると言った具合である。評価の世界では、この枠組を「ロジック・モデル」(Owen 2006)、「プログラム理論」(Rogers et al. 2006; Rossi et al. 2004)、「変化の理論」(Patton 1997)と多様に表現する。ここでは、JICA も含めて多くの援助機関や国際機関が使っているロジック・モデルに的を絞って、簡単な説明を加えておこう(国際協力機構 2004b)。

ロジック・モデルは、特定の課題に対する取組が理論上どう組み立てられ、最終成果がどのように実現するかを簡潔に示すためのツールで、取組・活動、必要な投入、活動結果、成果・インパクトをフローチャート、表、図等を使って体系的に表現する。取組から期待される成果・インパクトがこのプロジェクトで目指す目標に対応する。プロジェクトによっては、「Aを実施すればBが実現し、BのインパクトでCが達成される」というように、活動(A)の先に成果に対応する中間的な目標(B)とインパクトに相当する最終的な目標(C)が想定されることもある。その場合の因果関係の想定は、「目標の体系」として捉えられる。また、教育開発のように長期にわたる事業では、教育政策の変更のようにプロジェクトを取り囲む状況が変わることがよくある。ロジック・モデルでは、予めプロジェクトの理論が作用するた

めの前提条件の形で鍵となる外部的与件も特定する。

パネル1は、仮説的な小学校施設整備プロジェクトをロジック・モデルで表現したものである。事業でロジック・モデルを活用することの価値は、その作成プロセスにある。因果関係の想定や外部条件の特定をし、モデルの素案を作ることは、経験豊富な事業管理者であればそれほど難しくないかもしれない。しかし実施段階で多様な形で係わる関係者がその想定通りに行動するか、そして行動の結果期待される効果が生まれるかは不確かな場合の方が多い。できるだけ作成プロセスで利害関係者の意見を聞き、分析を重ね、それに基づく判断や見通しを取組の構築に反映する方が実現性の高いモデルとなることは明らかである。そうすることで、取組のあり方と期待される成果について関係者と共通理解を持つことができ、事業の実施に筋道をつけることができる。このように、ロジック・モデルの活用は、事業実施に理論的枠組みを付与するだけでなく、評価についても理論と実践の橋渡しをしてくれる。

ロジック・モデルの活用でよくあるもうひとつのケースは、事業が終了し評価を実施する段階になって取組と成果を結ぶ明確な因果の想定が不明瞭なことが判明する場合である。大規模で多様な介入を組み合わせた事業では、事業関係者の間で事業の組み立てについて共通理解を持たないままに活動だけが進んでいってしまうことが多々ある。そのような場合には、評価を行うために事業活動の体系を見直し、成果目標に応じてロジック・フレームを事後的に作成することが試みられる。国際開発高等教育機構がPCM手法を活用して作成したLEAD評価の手法はこの種の試みに有用で、それも理論と実践を橋渡しするツールと言ってよい(国際開発高等教育機構 2003)。

パネル1 小学校施設整備プロジェクトのロジック・モデル

課題：	学校施設の整備上の欠陥（教室数の不足、水供給の不備、トイレ施設の未整備等）のため、児童・保護者に対して登校に対する誘引動機を提供することができず、結果として低就学の状況が現出している。
活動：	①地域・保護者に対する小学校施設整備計画の説明 ②小学校施設の整備（教室の追加、水道の整備、トイレの新設等） ③学校設備メンテナンス体制の確立(含研修)
投入：	・追加・新規施設建設のための技術的・人的・資材的・資金的インプット ・メンテナンス専門家 ・教職員の時間
結果：	・整備済み施設（教室数の追加、水道施設、新設トイレ） ・メンテナンス研修を受講した教職員、児童・生徒
成果：	・在校生の出席率の改善 ・新規登録児童・生徒の増加による地域内就学率の向上 ・改善施設の維持
外部条件：	教室数の増加に見合う教員数の増加が実現することにより、対生徒比率が悪化することがなく、教育の質が維持される

4 . 教育援助分野の事後評価の実践：世界銀行の場合

事後評価の実施方法は、各援助機関によって異なる。ここでは、組織的な取組として最も進化していると考えられる世界銀行を例にとって考えてみることにする⁽²⁾。

世銀の事業評価慣行の特徴は、事業担当当局による自己評価を基軸とする重層的な内部評価の仕組みである。世銀は銀行組織であるから、事業立ち上げ時の事前投資案件調査が徹底していることはもちろんであるが、それに加えて事業の品質管理グループ（Quality Assurance Group）による事前段階でのチェックを行っている。このチェックはランダム選択ベースで、全案件数の1/4程度が対象となるが、平均的な事業の質の判定に役立っている。年によって違うが、最近年で92%が「良好」あるいは「ほぼ良好」の範疇に入っているようである（World Bank 2006）。この事前チェックの持つ意味は、新規案件の事前評価が担当部局外の専門単位に

よって行われることにあり、選択ベースでありながらかなり全案件に対して案件デザインの良質化のインパクトを持つと考えられている。

世銀の事後評価慣行の基本は、全ての事業の終了時（融資終了日から6ヶ月以内）に、担当部局（地域部）が作成する「実施完了報告（Implementation Completion Report、以下ではICRと略す）」である。ICR作成の目的は、事業に関する経験・知識・教訓の共有とアカウントビリティ（説明責任）で、その作成様式は標準化されていてコンピュータ画面上で作業するようにマニュアルができており、重要項目については「ベスト・プラクティス」事例がホームページで提供されている。ICRの構成は、(a)事業概要データ、(b)主要実績評定、(c)開発目標、デザインおよび初期条件のアセスメント、(d)目標とアウトプット達成度、(e)実施と成果に対する主要な影響要因、(f)自立発展性、(g)世銀および融資先のパフォーマンス、(h)教訓、(i)パートナーの見解、(j)追加的情報、(k)添付資料の12項目

で、添付資料を除いて通常シングル・スペースで7頁以内(複雑な事業では10頁以内)とされている。最も重要な(b)の主要実績判定は、下記の通り成果が6段階評価、他は4段階評価される。

- ・ 成果： かなり満足 / 満足 / わずかに満足 / わずかに不満足 / 不満足 / かなり不満足
- ・ 自立発展性： かなり見込みあり / 見込みあり / 見込みなし / とうてい見込みなし
- ・ 組織制度的開発インパクト： 高い / かなりあり / 低い / ほとんどなし
- ・ 世銀パフォーマンス： かなり満足 / 満足 / 不満足 / かなり不満足
- ・ 融資先パフォーマンス： かなり満足 / 満足 / 不満足 / かなり不満足

判定の基準については、ガイドラインで説明が付与されているが、判定の根拠は次の諸項目から成る添付資料で提供されるようになっていく。

主要実績指標(ログ・フレーム・マトリックスの提示)

事業コスト(活動項目別; 拠出先別; 支出費目別)

費用便益分析

世銀インプット(現場訪問、スタッフ投入)

目標達成度の評定

世銀と融資先のパフォーマンスの評定

補足文書

世界銀行の教育分野における融資事業は、1件当たりの金額が5,000万ドルから2億ドルの金額が示すとおり、かなり規模が大きい。従って、個々の事業を構成する要素もたくさんあり、事業の成果、自立発展性、組織制度的開発インパクト等について評価するための指標も多種多数になる。例えば、1990年代後半に実施されたガーナの基礎教育セクター改善事業(World Bank 2003)の場合、(i)

教育・学習の質の改善、(ii)経営管理の改善、(iii)教育アクセスの改善の3事業要素で構成されているが、ロジカル・フレームワークの成果目標として、(i)については学力テスト結果、教員の欠勤率、主要科目の生徒・教科書比率等4指標、(ii)については生徒・教員比率、教員・管理者比率等3指標、(iii)については学年別就学率、留年率、修了比率等8指標、計15指標が挙げられた。さらに財政指標として、政府予算中の教育予算比率、教育予算中の基礎教育支出比率、給与外支出比率等も加えられている。必ずしも全ての指標について、事業終了時の到達目標値を特定しているわけではないが、ICRではそれぞれについて事業開始年の当初値と終了時の実績値を提示している。費用便益分析については、融資決定の一助としてかなり無理をして初中等教育の内部収益率を計算しているが、ICRでは教育投資の収益を計算するには早過ぎるとして数値をはじいていない。他の教育事業についてもほぼ同様のようで、初中等教育に対する支援の場合には、基本的に事前・事後の就学率と学力テスト結果の変化を主とし、教育組織の管理効率の改善を従として、諸目標の判定をしており、経済的・財務的収益計算はしていないようである。

世銀には内部に独立した「評価者」の役割を持つ業務評価局(Operations Evaluation Department、以下OEDと略す)がある。理事会に直属する独立部局で、評価手法の開発も含めて評価活動全般に責任を持つ。OEDは、事業終了時に各部局が作成するICRの質を主にデスク・ワークでレビューし、評価要約(Evaluation Summary、以下ESと略す)をまとめる。これは担当部局評価の2次評価に当たり、その際にICRの成果、自立可能性等の判定を是認することもあるし、修正を加えることもある。OEDの刊行する年次報告書によると、レビューしたICRの92%は、総合的な質の評価で「満足」以上のレベルにあったとしている。しかし、ICRの事後経済

分析については、「満足」以上のレベルにあるのは85%、判定の検証材料が十分で説得的なのは83%とかなり辛口の評価もしている。事業の将来的なモニター・評価に関する計画については75%で、自立発展性について世銀プロジェクトでも問題があることを示唆している。

OEDは、さらに、全終了事業の約1/4の案件について独自の評価調査を行い、事業パフォーマンス評価レポート（Project Performance Assessment Report、以下PPARと略す）を作成する。これは限定案件に関する3次評価に相当する。PPAR調査では既出資料に加えて、通常1週間程度の現地出張による事業関係者へのインタビュー等により定性的インプットを加えてより厳密な評価を下す。世銀の内部評価メカニズムがどの程度うまく機能しているかを測る指標として、ICRの成果評価の結果がES、PPARの2次、3次評価の段階でどう修正されるかをみることができる。例えば、1998年度と2000年度には担当部局の成果の判定に対するOEDの2次評価（ICR ES）で4～5%（20～25件に1件）が「満足」段階から「不満足」段階に下方修正されたが、上方修正の件数は比較的少なかった。OED内での再評価に当たる3次評価（ES PPAR）では、2年度ともにESの評価をさらに8～10%（10～12件に1件）下方修正したが、ほぼ同等の比率で上方修正も加えた。両者を合わせると、15～20%（5～7件に1件）が判定を修正される計算である。要するに、事業担当者の評価判定は、OEDの厳密評価よりかなり「甘い」確率が高いということで、OED当局は判定の再検証を2次評価段階でとどめず、より多い案件について3次評価を実施すべきとの議論が成立する。

以上の世銀の内部評価システムが教育援助事業の評価に実際にどう適用されるかについてブラジルに対する基礎教育援助事業のPPAR（World Bank 2002）を例に引いて

みてみよう。このPPARでは、1990年代に実施された4件の関連融資プロジェクトを評価している。それぞれの案件につき、成果、自立発展性、組織制度的開発、融資先パフォーマンス、世銀パフォーマンスの5項目について、ICR、ES、PPARと3段階の評定を比較しているが、パネル2は、項目ごとに4件の2次、3次評価による判定修正がどうであったかを示したものである。

この事例の4案件の場合には、2次評価（ICR ES）の段階で、成果だけでなく、組織制度的開発、世銀パフォーマンスについても3件で下方修正が行われている。3次評価の段階では、さらなる下方修正はないが、成果目標に関する判定では下方修正された3件のうち2件について上方修正している。これは、細かくみると2件ともICRで「満足」の判定をESが「わずかに不満足」と判定したものを、「わずかに満足」とやや戻したものを、それでもOEDの判定は事業担当部門の判定よりは厳しく出ていることがわかる。この例示案件の場合、OED内の修正はあまり多くない。教育援助事業の場合、効果の発現の把握が難しいだけに、PPARで現地調査してもデスク判定を大幅に覆す判定には導かないことを示唆しているのかもしれない。パネル2で興味深いのは、融資先パフォーマンスについて、概ねICRの判断を踏襲しているのに対して、世銀パフォーマンスに対してICRの判定を下方修正していることで、事業担当部局に対するOEDの独立性を象徴している。

世銀のICRを起点とする評価システムは、主にアカウンタビリティを目的としているが、OEDの機能の中には実施事業からの経験と教訓のフィードバックが含まれている。実際、ICRでもかなりの頁を割いているが、2000年から従来のICRを「中核アカウンタビリティICR(Core Accountability ICR)」とした上で、主に事業経験からの教訓導出に的を絞った「集中的学習ICR(Intensive Learning ICR)」の範疇を新たに設けてい

パネル2 世銀の教育援助事業の2次、3次評価による判定修正
(ブラジルの基礎教育支援プロジェクトの4案件の例)

判定項目	ICR→ES			ES→PPAR		
	下方	上方	無修正	下方	上方	無修正
成果	3	0	1	0	2	2
自立発展性	0	0	4	0	1	3
組織制度的開発	3	0	1	0	0	4
融資先パフォーマンス	1	0	3	0	0	4
世銀パフォーマンス	3	0	1	0	1	3

(出所) World Bank (2002, p. v)の判定総括表より作成

る。ILICR案件は全案件の約30%の見当で、フォローアップの可能性のある事業、パイロット事業、新規事業主体、問題発生事業等の基準に基づき、組織全体の学習の観点から選ばれることになっている。CAICRと比較すると、ILICRにはより詳しい説明が加わる(但し、使う様式は同じ)ほか、作成の過程で、事業からの教訓に関する利害関係者ワークショップを開くこと、それに向けて受益者調査を行うこと、他の関連部局(地域)スタッフが学習目的で参加すること、OEDが場合によっては並行的に監査を実施し、担当部局の評価をレビューすること等々の追加要件が見つく。

5. 教育援助評価の実践的課題

以上見てきたように、教育援助評価の事後評価については、ロジック・モデルの適用や援助評価システムの進化により、援助関係者間での共通理解は進んできた。各国援助機関および国際機関の共同による基礎教育分野の援助のインパクト評価の試み(Netherlands Ministry of Foreign Affairs 2003)は、そのことを示唆している。それは同時に評価の理論とそれに基づく評価手法の定着を示唆している。しかし、評価の実践レベルでは、それだけでは解決できない重要な課題がいくつか

ある。

第1の課題は、理論的配慮をいかに評価の実践に反映するかで、特に評価において不可欠な価値判断をどう事業の理論的枠組と矛盾しない形で取り込むかである。例えば、教育アクセスの拡大と教育の質の向上というように場合によっては二律背反的な(しかし二者択一できない)ふたつの政策目標が掲げられている時に、いかに評価の基準を設けるかといった課題が生じる。希少な学校建設資源を、教育アクセスの限られている農村に向けるか、児童数が多く教室が過密状態で教育の質の劣悪な都市に向けるかで、実際の政策が明確な配分基準のないままに実施された時に、政策インパクトをどう評価するかといった課題に理論はどう答えてくれるのだろうか。この関連で注目に値するのは、米国の評価専門家間で普及しているチェック・リスト・アプローチで、評価の理論と実践のバランスを標準的な評価手順に置き換えて、簡便な評価を提案している(Davidson 2006)。例えば、複数の事業目的間で相対的な重要性の判定をしなければならない時に、利害関係者の投票、専門家の判断、当該分野の文献の活用等により目的間のウェイト付けを行う手法である。それは、手法の透明性を確保することにより、実践的な困難を緩和する試みである。

第2の課題は、開発途上国政府の教育開発政策との整合性をどう図るかである。EFAの推進のためのFast Track Initiativeでは、世銀を中心とするドナー・コミュニティの協調支援の枠の中で途上国政府が「運転席」に座り、セクター・ワイド・アプローチを推進しようとしている。援助活動の評価の軸となる成果の設定は当然途上国の教育政策に基づくものでなければならないが、それが援助国側の目標設定と整合的である保証は必ずしもない。また、伝統的なプロジェクト・ベースの援助と異なり、多数の援助国による多様な取組で取組と成果間の因果関係が複雑であること、特定の介入のインパクトの効果を把握することが難しいこと等のため、個々の援助国にとって「出資相当分」の評価を行うことは極めて困難となる。この実践的な課題は、各ドナー国における評価文化の浸透を待つより仕方がないのかもしれない。

第3の課題は、教育援助事業の持続可能性の評価をどうするかである。持続可能性はDAC評価5基準のひとつであるが、その評価方法は確立されていない。通常の終了時評価では、援助受入国側の政府や組織が援助終了後もプロジェクト活動を継続するための予算的・人材的措置を講じているかをチェックする程度である。教育の効果の発現には長期を要することから、教育援助は短期的な効果を狙うより、教育システムの構築・整備に向けられるべきとの考え方がある。その場合には評価の主要観点が持続可能性の検証になるが、住民参加といった受入国側の参加を重視するそのようなプロジェクトでは、支援対象となる組織とそれをめぐる状況に関する政治・経済・社会的分析を必要とすることもあり、画一的に考えることはできない(Platteau 2004)。

最後に、第4の課題として、援助評価の国際化がある。援助の有効性については、常に援助懐疑論者からの批判が付きまとい、「厳密な」インパクト評価の実施が叫ばれている

(Center for Global Development 2006)。ODA評価にまつわるあいまいさを除くには、無作為サンプル方式を活用した評価を多用すべきとの考え方は、世銀内部でも広がっており、二国間ドナー機関も無視できなくなってきた(Pitman et al. 2004, Part 3)。EFAの動きの中で評価活動についても国際的に共通理解が進み、共同の取組が普及した結果として、評価の実践についても各国が独自の評価システムや慣行を維持することが難しくなってきた。日本のようにプロジェクト・ベースの援助パターンを維持している国にとっては、援助評価の国際化に折り合いをつけることは必ずしも容易ではない。しかし、国際社会で日本の援助政策や援助方法について正しい理解を広げるには、単に評価に関する国際的議論に参加するだけでなく、積極的に自らの評価慣行について発言することが必要であろう。

6. 結びに代えて

本研究ノートでは、教育援助プロジェクトの事後評価について、理論的裏づけと実践的課題について整理した。事後評価については主に終了時評価を念頭においてまとめたが、プロジェクト完了後3～5年の期間において主としてインパクトに的を絞っておこなう事後評価についてもほぼ本稿で書いたことは当てはまる。援助機関が中心になって行う援助評価では、常に評価手法や基準の一貫性と評価対象となる個々のプロジェクトの固有の状況の間で適正なバランスをとることが求められる。この課題に対する解は常に実践における反省の形で与えられるが、そこにおける経験則はプロジェクト開始前に十分な事前評価を行うことによりプロジェクトの設計段階でその実施と効果の現出について確度の高い見当をつけることにある。評価理論の理解と応用の深化がそのプロセスを支えていくことは言うまでもない。

注

- (1) 本稿は、平成18年度に広島大学が国際協力銀行から受託して実施した「教育ネットワーク研究会 国際教育開発連続講座」で筆者が使用した講義テキストの一部に加筆修正したものである。
- (2) 以下の記述は長尾（2003）に依拠している。

参考文献

国際開発高等教育機構（2003）『開発援助のためのプロジェクト・サイクル・マネジメント：モニタリング・評価編』国際開発高等教育機構・国際協力機構（2004a）『評価結果の総合分析：初中等教育 / 理数科分野』国際協力機構。

国際協力機構（2004b）『プロジェクト評価の実践的手法』国際協力機構。

国際協力機構（2006）『評価結果の総合分析：高等教育』国際協力機構。

長尾眞文（2003）「教育援助評価の現状と課題」『国際教育協力論集』6巻1号、1-18頁。

Campbell, D. T. & Russo, M. J. (1999). *Social Experimentation*. Thousand Oaks: Sage Publications.

Center for Global Development (2006). *When Will We Ever Learn? Improving Lives Through Impact Evaluation*. Report of the Evaluation Gap Working Group (May 2006), Washington, D. C.

Chapman, D. & Nagao, M. (2006). Editorial: Assuring Quality and Effectiveness during Rapid Change. *Journal of International Cooperation in Education*, Special Issue: The Evaluation of International Cooperation in Education, 9(1), pp.1-5.

Davidson, J. (2005). *Evaluation Methodology Basics: The Nuts and Bolts of Sound Evaluation*. Thousand Oaks, CA: Sage Publications.

Netherlands Ministry of Foreign Affairs (2003). *Local Solutions to Global Challenges: Towards Effective Partnership in Basic Education*, Joint Evaluation of External Support to Basic Education in Developing Countries, The Hague.

Owen, J. (2006). *Program Evaluation: Forms and Approaches*. Crows Nest NSW: Allen & Unwin.

Patton, M. (1997) *Utilization-Focused Evaluation: The New Century Text*. 3rd Edition. Thousand Oaks, CA: Sage Publications. (日本語抄訳版: マイケル・クイン・パットン著 (2001) 『実用重視の事業評価入門』大森彌監修、山本泰・長尾眞文編, 清水弘文堂書房)

Pitman, G. K., Feinstein, O. N. & Ingram, G. K. (2004) *Evaluating Development Effectiveness, World Bank Series on Evaluation and Development*. Vol. 7, New Brunswick & London: Transaction Publishers.

Platteau, J. P. (2005). Institutional and Distributional Aspects of Sustainability in Community-Driven Development. In Pitman, G. K., Feinstein, O. N. & Ingram, G. K. (eds.), *Evaluating Development Effectiveness, World Bank Series on Evaluation and Development*, Vol. 7, New Brunswick, N. J.: Transaction Publishers (pp. 275-297).

Rogers, P. J., Hacsı, T. A., Petrosino, A. & Huebner, E. A. (eds.) (2000). *Program Theory in Evaluation: Challenges and Opportunities*. New Directions in Evaluation, No. 87, San Francisco: Jossey-Bass.

Rossi, P. H., Lipsey, M. W. & Freeman, H. E. (2004). *Evaluation: A Systematic Approach*, 7th Edition, Thousand Oaks, CA. (日本語版: ロッシほか (2005) 『プログラム評価の理論と方法』大島巖他監訳, 日本評論社)

Shadish, W. R., Cook, T. D. & Leviton, L. C. (1991). *Foundations of Program Evaluation: Theories of Practice*. Newbury Park, CA: Sage Publications.

World Bank (2002). *Project Performance Assessment Report: Brazil (Innovation in Basic Education Project, etc.)*. Washington, D. C.: The World Bank.

World Bank (2003). *Implementation Complete Report on a Credit to the Republic of Ghana for a Basic Education Sector Improvement Project*. Report No: 26095, Washington, D. C.

World Bank (2006). *Quality Assurance Group: Assessment of the Quality of "Other Economic and Sector Work"* (March 15, 2005). <<http://siterources.worldbank.org/QAG/Resources/OESWSynthesisReport.pdf>>