

教育援助評価の現状と課題

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

援助、特に政府開発援助(ODA)の評価の慣行が本格化したのは、1982年にOECDの開発援助委員会(Development Assistance Committee、以下ではDACと略す)の下に援助評価に関する専門部会が設けられてからである。「評価グループ」と呼ばれるこの専門部会は主として各国援助機関の評価担当者で構成されており、主たる活動内容として評価手法や判断基準に関する援助国間の共通理解の形成や合同評価の実施、さらには発展途上国の評価能力の育成を行ってきた。その貢献として特記されるのは、いわゆるDAC評価5基準 - 「妥当性(relevance)」、「有効性(effectiveness)」、「効率性(efficiency)」、「インパクト(impact)」、「自立発展性(sustainability)」 - を設定したことである。それにより援助分野の別なく、全ての二国間・多国間援助機関がほぼ共通の評価基準を採用するようになり、援助慣行の定着を助長した⁽¹⁾。しかし、1990年代に先進諸国ではほぼ例外なく実施された公共部門の行財政改革は、政府開発援助の分野でも成果重視の傾向を助長し、各国の援助機関に対してODA評価の効果的・効率的実施についてそれ以前よりも厳しくアカウンタビリティ(説明責任)を問うようになった(DAC 2000、p.3)。実際にDAC諸国のODA総額が1990年代の10年間に実質タームで60億ドルから50億ドルへと16%強減少したことも手伝って、援助事業の見直しや評価に対する関心が高まった。

教育援助は、最も古くからの援助分野のひとつであり、援助国の2国間援助総額の8~

10%(UNESCO 2002)を占める重要分野である。援助に対する一般的な引き締めが強化される中で、教育分野の援助は内容が納税者に比較的理解されやすいこともあって、アカウンタビリティのための評価の対象として相応の扱いを受けてきた。教育援助評価が特に注目を集めるようになったのは、1990年代に国際教育協力の分野でも新たな展開が起きたからである。1990年にタイのジョムティエン(Jomtien)で開かれた「万人のための教育(Education for All/以下EFA)」世界会議は、基礎教育の普及を途上国の教育開発の主要課題として位置づけるとともに、それまで援助国が個々に行っていた教育援助を国際教育協力の同一の俎上におき、援助国横並びで議論をするきっかけを作った。ジョムティエン会議はまた、10年後の2000年に教育普及に関するEFA目標の達成について検証することに合意することで、教育援助評価を国際教育協力の中心的な課題に据えた。2000年4月にセネガルのダカールで開かれたEFA世界会議では、各国、各地域による10年間のEFA目標達成努力とその成果が報告され、EFAプロセスの総合的な評価が行われた(International Consultative Forum on Education for All 2000)。筆者は、本論集第2巻第2号(長尾1999、148頁)において、この「万人のための教育」の評価が国際教育協力分野における評価慣行の定着に結びつくのかとの問題提起を行った。本稿では、その後の教育援助評価の研究と実践の動向を概観するとともに、そこでの主要な課題について述べることにする。日本の教育援助の評価については、牟田(1998)が1885年度 - 1997年度の期間に外務省により実施された

63 の教育案件事後評価の総括的検討に基づく経験と課題の整理を行っており、本稿ではより国際的な視点から課題の整理を試みる。

以下では、まず第2節で発展途上国に対する援助の評価に関する最近の動向について簡単にレビューした後に、第3節で教育援助評価の考察の概念的背景として米国における教育評価の発達とその国際的普及、特に発展途上国における教育評価実施上の問題について考察する。続く第4節で個別教育援助事業の評価の現状をレビューした後、第5節で教育分野のプログラム・政策支援の評価へのシフト、評価における国際連携等の新しい評価課題を取り上げる。最終第6節では、教育援助評価の今後の展開について述べることにする。

2. 援助評価の最近の動向

教育援助評価は、教育分野における援助事業の評価のことであるが、そこにおける評価のあり方は、当然一般的な援助評価のあり方や動向に左右される。そこで教育援助評価の現状と課題に入る前に、援助評価の一般的な動向についてレビューしておきたい。日本評価学会の学会誌『日本評価研究』は2003年3月に刊行した第3巻第1号で「政府開発援助（ODA）における評価の新たな潮流」のテーマで特集を組んでいるが、その冒頭で編者の三好（2003）は、ODA評価の現状の整理・再検討の主要な観点として次の3点を挙げている。第1は、援助の成果重視の志向により、開発協力の有効性に加えて途上国の開発の有効性の視点を強調する必要性が生じ、それが評価実施体制の再検討を迫っているとの認識。第2は、援助評価の対象を従来からの事業（プロジェクト）一辺倒からそれらを束ねたプログラム、さらにはそれらを包含する政策にまで拡大する必要があるとの認識。そして第3は、事後評価に集中する評価範囲を改めて、事前、期中も加えた事業サイクル全

体の評価を実施する必要があるとの認識である。

従来の援助評価は、個別援助事業の実施に関するモニタ - とせいぜいアウトプット（産出物）の確認までであり、実際に事業の目的や目標がどこまで達成されたのか、想定受益集団に対するインパクトがどれだけあったのかについて系統的に検証することに注意が向かなかった（Bamberger 2000, p.96; DAC 2000, p.18）。前述したように、1990年代に進行した援助の成果重視の傾向が、評価により積極的な役割を付与する方向に変化させたのである。第1点の途上国の開発へのインパクト重視の傾向は、DACが1998年に実施した開発援助評価のDAC原則の見直し（DAC 1998）で取り上げられており、スウェーデン（Svensson 1997）、日本（外務省 2002）等でも明確な姿勢を打ち出している。第2の援助評価の対象を事業だけからプログラム、政策も含む方向へ転換する動きは、1980年代から始まった構造調整支援、1990年代の包括的開発フレームワークや貧困削減戦略の推進と政策レベル援助傾向を強化した世界銀行（以下では世銀と略す）が急先鋒となって動いてきた（Picciotto 1997）。この傾向は米国をはじめいくつかの援助国でも見られ、ジェンダーや環境といった部門にまたがる援助の展開の形としても実現しつつある（DAC 1998, p.7）。第3の事業サイクル全体を評価する傾向は、民間企業における成果重視マネジメントの考え方を援助分野に導入したことの当然の帰結である。日本でも国際協力事業団（JICA）が明確にこの方向に動いており（国際協力事業団 2001）、外務省、国際協力銀行でも事前評価手法に関する調査等に着手している（外務省 2003、国際協力銀行 2002）。

以上にも関連するが、援助事業の評価プロセスに事業受益者も含めて被援助国側の主体が参加する参加型評価の実践も目立つようになってきた（Bamberger 2000, p.97）。それ

は先進国援助機関にとって評価の質の改善と一般国民に対する評価の信頼性の補強を同時に可能にする手段となる可能性を持っている。また、1990年代の援助事業をめぐる議論の基本的なテーマであった発展途上国の当事者意識を助長する効果も期待できる。その関連で、被援助国側の評価能力の強化も注目を集めた(DAC1998, p. 11)。しかし、援助評価に直接関与してそこから収入も得る途上国専門家は別として、被援助国側には援助国側と同様の援助評価の実施に対する動機は存在しないし、評価の概念さえ一般的に理解されているとは言えない(Garaway 2003, p.708)。援助国側と被援助国側の評価関心には顕著な非対称性がある(長尾 2001, p. 90)。発展途上国の論者の中には、自国の公共部門の改革に評価が戦略的な重要性を持つとの意見を述べる者もいる(Wiesner 1997)。しかし、そのような改革と無縁の多くの途上国で、しかも援助事業の実施が参加型になっていない時に、評価プロセスだけを参加型にしても被援助国の積極的関与はあまり期待できないであろう。

3. 教育評価の発達と国際的普及

教育評価は⁽²⁾、主に米国で発達してきた実践的科学であるが⁽³⁾、その源泉は学校における生徒児童の学力測定(アセスメント)と学校の認証のための教育プログラムの質の検証に求められる(Kellaghan et. al. 2003, p.2)。前者は、学力テストによる到達度の把握が主たる内容で、20世紀初頭から教育学者により定量的な手法で学力水準の測定の厳密性を問う試みが繰り返されてきたが、教育内容の質的評価は必ずしも対象とされていなかった。教育評価が教育カリキュラムや教育活動の内容の評価に踏み込むようになった契機は、1930年代にRalph Tylerが行った高校教育が大学生の学業に及ぼす影響の「8年間調査」であった。Tylerは、高校教育の中味

が大学に進学した学生の学業進度に深くかわることを証明することにより、教育目的に従って教育カリキュラムや方法を評価することの意義を唱え、今日の教育評価の素地を築いた(Popham 1993)。後者の学校認証制度のための小中高校および大学等教育機関の評価は、内部評価と外部評価を組み合わせることで学校経営の継続的改善を目指すもので、米国では公的な教育制度を補完する慣行として定着している。この双方の伝統が相俟って教育評価の分野を形成するようになったのは、1960年代中ごろから1970年代にかけて米国連邦政府が教育の質の向上を目的として各種の教育事業に巨額の投資をするようになり、その有効性や効率性を追跡するために評価活動を奨励するようになってからである。その後の展開で、テスト重視の定量的調査志向は、教育事業や教育機関・システムの産み出す成果の検証に向けられるようになり、多様な教育評価が実施されるようになった。また、教育評価は、他の社会的・人的サービス分野における評価の普及にも、多大の影響を与えた⁽⁴⁾。

Kellaghan and Stufflebeam(2003)によると、教育評価は、3つの点で他の分野の評価と顕著に異なる。第1点は、前述した学力測定(アセスメント)の伝統が未だ残っていて、事業評価と関連するが別の領域として「教育アセスメント・テスト」を抱えていること。第2は、教育は他の分野と違って、社会のほぼ全ての構成員に何らかの形で関係していることから、評価を実施する際に利害関係者の意見の汲み上げにかなり配慮する必要があること。そして、第3点として、教育評価では、教員が、ある時は評価者として、ある時は評価される対象として、またある時は利害関係者として関与するが、数が多いだけでなく、教育活動の中核を占めるので、評価を実施する際に無視できないこと、以上3点である。後の2点は、評価の実施プロセスを複雑にするが、Psacharopoulos(1995)は、

教育事業の評価を難しくする点として、教育効果の発現に長時間を要すること、教育内容が多様な上にさらに複雑化する傾向にあること、教育関心が多元化し判断が難しくなっていること等を挙げる。

教育評価の慣行は、北米を起点として欧州諸国、日本をはじめ他の地域に普及した⁽⁵⁾。当然、国により地域によって教育評価を実施する社会的、文化的な文脈は異なり、政治的な状況も同様ではない。そこで生じる多様な問題について、Kellaghan and Stufflebeam (2003) 編による *International Handbook of Educational Evaluation* (教育評価の国際ハンドブック) では、「教育評価の社会的、文化的文脈」と題するセクションを設けて、米国、欧州、ラテン・アメリカ、アフリカの4地域の視点から論じている。Reimers (2003, p.461) は、ラテン・アメリカにおける教育評価の慣行について、国家の統制とエリート優先に特徴付けられる教育政策の制度的な枠組みに支配されており、評価の実施に係わる国際機関や国内研究機関の専門家もその中に取り込まれているとする。また Omolewa and Kellaghan (2003, pp. 478-479) は、アフリカにおける状況について、部族社会の伝統の中に人材育成に形成的評価を活用する慣習があるが、国家主導による近代的教育普及の動きに押されて忘れられており、輸入慣行としての児童生徒の学力評価、教育システムの評価、教育事業の評価のどれをとっても動員できる資源の不足により、十分な成果を挙げていないと論じている。前述した教育事業の評価で不可避な難しさに加えて、発展途上地域にはそれぞれに特有な社会的・文化的文脈があることから、教育評価を定型的に考えることは難しいと言える。

しかし、冒頭で触れた DAC の評価基準の国際的共有化と同様に、教育評価に関する考え方についても、「グローバル化」が緩やかに進行しつつあると、Bhola (2003) は議論する。国境を越えた教育専門家の共同研究や

学会参加を通じる国際的な交流を通して、非西欧世界の評価者や評価研究者も西欧発の評価の基準を受容しつつあるとの見地である。その典型的な例は、UNESCO の教育計画に関する研修機関である International Institute of Educational Planning (IIEP) が南部アフリカ7カ国で実施している初等教育の質の改善に関する共同研究事業である⁽⁶⁾。この事業では、参加各国政府の教育政策担当者の積極的な関与のもとで、共同で実施する学校レベルの研究手法について研修することから出発して、研究計画の並行的な立案、実施、さらには研究成果の交換へと展開しており、教育評価手法の国際的な普及の例として高く評価されている (Postlethwaite 1999)。また、近年特に顕著になった評価慣行の国際的普及要因として、各国や地域で設立されている評価学会の活動がある。例えば、アフリカ地域では各国の評価者が国際機関の評価者等の支援を受けて、アフリカ評価学会を設立しているが、その活動のひとつとして、米国・カナダの評価基準にならって「アフリカの評価基準」を定めようとしている⁽⁷⁾。また2003年3月には、40カ国・地域の評価学会がペルーのリマで International Organization for Cooperation in Evaluation という国際的な評価学会の連携組織を結成した⁽⁸⁾。その標榜するところは、各国・地域それぞれの社会的、文化的文脈の特異性を重視する評価の実践であるが、このような動きも長期的には評価一般、ひいては教育評価のグローバル化を後押しするものとなる。

4. 個別教育援助事業の評価の現状

筆者は前掲論文 (長尾 1999) で、個別教育援助事業の評価に係わる研究課題として、「評価の実践において評価目的が生かされているか」、「事業内容に即した評価が行われているか」、「被援助国と評価について共通理解

があるか」の3点を挙げた。これらは教育事業評価の手法的あるいは理論的関心に沿う設問というより、援助評価システムのあり方をより重視する実践的設問で、実際の援助評価作業の経験的分析の必要を示唆している(長尾 2001)⁽⁹⁾。教育分野に限らず一般的に個別援助事業の評価を最も系統的に実施している援助機関は世銀である。以下では、まず世銀の個別教育援助事業の評価に係わるシステムと慣行について詳しく検討し、次にそれとの対比で一部先進国援助機関の取組みについて触れることにする。

(1) 世界銀行による教育援助評価

世銀の事業評価慣行の特徴は、事業担当部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⁽¹⁰⁾。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)主要実績評価、(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⁽¹¹⁾、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

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)

目標達成度の評価

世銀と融資先のパフォーマンスの評価
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。⁽¹²⁾ 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(Effron 2002, Table A1.2)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。表1は、この重層的評価の進行を、各段階を経た件数のうち「満足」のつく評価から「不満足」のつく評価に下方修正された件数の比率とその反対に上方修正された件数の比率で表わしたものである⁽¹³⁾。

表1 成果評価の2次、3次評価での判定修正案件の比率
(検討案件に対する比率：%)

評価段階	1998年度				2000年度			
	下方	上方	合計	ネット	下方	上方	合計	ネット
2次：ICR→ES	4	1	5	3	5	0	5	5
3次：ES→PPAR	8	7	15	1	10	10	20	0

出所： Effron (2002, Table A1.3)より作成。

この表を見ると、1998年度、2000年度ともに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ていることが分かる。上方修正の件数は比較的少ない。次にOED内での再評価に当たる3次評価(ES PPAR)を見ると、2年度ともにESの評価をさらに8～10%(10～12件に1件)下方修正しているが、ほぼ同等の比率で上方修正も加えている。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ある。以上から、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を導こうとしている⁽¹⁴⁾。

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(世界銀行2002)を例に引いてみよう。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、表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。但し、表1と異なり、表2では上方・下方修正を「満足」段階と「不満足」段階の2段階ではなく、成果は6段階、それ以外の項目は4段階で修正の方向だけを見た。

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。表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。

表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。

世銀のICRを起点とする評価システムは、主にアカウント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「中核アカウントビリティ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つく。しかし、Efron(2002, pp.15 - 16)によると、ILICRの実施は、2000年の導入後1年余の期間で、未だ全体の10%程度に過ぎず、質的にも通常のCAICRと大差ない。「知識の銀行」を自称する世銀の内部評価システムでも、教訓導出の方向への本格的な稼働は、未だ先のようなのである。

(2) 先進国援助機関による教育援助評価

英国の政府援助機関(Department for International Development、以下DFIDと略す)も世銀と同様に独立した評価局(Evaluation Department、以下EvDと略す)を持ち、事業担当部局とともに実施事業の評価を行う。しかし、両者の評価作業は相互補完的ではあっても、世銀のように重層的・系統的に組まれているわけではない⁽¹⁵⁾。事業担当部局は、専ら実績管理目的のために

実施事業に関する評価を自らあるいは外部評価者に委託して実施する。DFID内の事業データベース用の事業要約もそれぞれで作成し、評価局が2次評価する構造にはなっていない。一方、評価局は議会に対するアカウントビリティと援助の有効性・効率性の改善を目的として事業案件を選んで独自の評価調査を行う。援助評価の実施方法については、評価局が評価ガイドラインを提示しているが、評価技法等については、評価を担当する専門家の責任範囲としている。但し、ロジカル・フレームワークの活用と、5段階の総合的評価 - 「かなり成功(Highly successful)」、「成功(Successful)」、「部分的に成功(Partially successful)」、「ほとんど不成功(Largely unsuccessful)」、「不成功(Unsuccessful)」 - による判定の慣行は定着しているようである。総合判定の方法としては、評価者が与えられたプロジェクト・パフォーマンス基準のリストの中から、評価対象事業に関連する基準を選び、その相対的重要性と成功度を評定をする。そこで特徴的なことは、成功度の判定を費用に対する目標達成の度合い、つまり費用対効果タームで考えていることである。

1997年に4ヶ月かけて実施されたケニアの初等教育強化プロジェクトの評価(DFID 1998)を例にとり、EvDがどのように機能するのか見よう。この事業は1991年～1996年に実施された事業で、現職教員支援センターの組織制度と人材の強化により小学校における理数科教育と英語教育の質を高めることが主たる目的であった。EvDの定義による評価目的は、その事業目的の達成に関する評価にとどまらず、就学率等の学校パフォーマンスへの影響、より幅広い社会的・経済的インパクトの検証、関連する政府とコミュニティの関係の問題の調査、さらにDFIDが実施中の調査への貢献等も加わり、文字通り総花的内容であった。評価チームは英国の教育コンサルタント団体で、EvDの

参加型評価の方針により、フィールドの調査には現地コンサルタントが協力した。この事業では出発時点でベースラインの調査を行っておらず、学校教育統計も不備なため、専ら学校関係者のインタビューや、フォーカス・グループといった定性的手法によるデータ収集が図られた。

評価の総合的判定は「部分的に成功」で、現職教員研修施設が増えて教員の研修機会が増え、教材開発が進んだこと等がプラス要因として、事業デザインの欠陥、受入国側の財政事情の悪化のインパクトとそのための研修施設の維持の困難等がマイナス要因として挙げられている。EvDの実施する評価は、世銀のPPARに相当するが、教訓導出の視点的に絞られていない。この評価を実施する段階で、既に同事業の次のフェーズが始まっていること、事業の直接関係者（受入れ国政府関係者とDFID担当部局）に対するフィードバックが組み込まれていないことで、評価結果がどのように使われるのか明らかでないことが指摘される。現地コンサルタントが評価作業を担当することで「参加型」にはなっているが、関係者の参加の評価に欠けている。

米国の政府援助機関（USAID）の場合には、評価関心が組織全体の実績管理に傾斜しており、全ての分野で個別事業の実施のモニタリングと終了時評価を担当部局の現場で分散管理している⁽¹⁶⁾。政策企画調整部局の中に設置されている開発情報評価センター（Center for Development Information and Evaluation、以下CDIEと略称する）が実施を支援する目的で評価実施の方法や成果指標について情報提供している（USAID 1996a、1996b）。USAIDは、ロジカル・フレームワークを用いて事業開始時に成果目標を可能な限り特定化・指標化する慣行の先駆的機関のひとつである。基本的に事業の実施プロセスを重視する姿勢で、期中管理を的確に行うことにより事業終了時の評価の負担を軽減する慣行である。CDIEでは独自に年間

100件程度の評価を行っているが、必ずしも個別事業案件を対象としておらず、むしろ優先援助分野での援助の有効性、効率性を検証する評価が主体である。教育分野はUSAIDの最重要視している分野のひとつで、評価に関する年次報告書(USAID 2003)でも1節を設けて論じている。評価作業の実施は、主として外部評価者への委託によることから、担当部局の評価契約管理責任者が重要な役割を果たす仕組みとなっている。

最後に日本の場合であるが、国際協力事業団では、JICA事業評価ガイドライン（JICA 2001）で通常の技術協力プロジェクトを念頭においた評価の実施方針と実施方法を提示している。事業開始当初にロジカル・フレームワークを活用して成果目標を規定し、期中、終了時とモニターしつつ評価していくアプローチである。しかし、成果その他の目標の立て方や判断については標準化しておらず、特に1990年代後半から実施が本格化した途上国の理数科教育支援等の教育分野における技術協力については、それぞれの事業で評価の取組みについて計画しているのが現状である。JICAが2001年度に設置した「JICA外部有識者評価委員会」では、試験的に終了時評価について2次評価を行う試みをおこなったが、それが世銀のような重層的、系統的な取組みに展開するかは定かではない。国際協力銀行では、円借款全案件を事後評価する体制が整い、事前評価の実施や第三者評価者の活用など社会サービス分野の評価活動の充実を図っているが（国際協力銀行 2002）教育分野における案件はまだ数が少ない。

5．新たな教育援助評価の課題

伝統的な個別教育援助事業に対する関心は、各援助機関が直面している状況に応じて変遷を迫られている。この状況変化に影響しているいくつかの要因は各国に共通であり、

その関連で新たな課題を産み出している。本節では、その中で国際教育協力に及ぼすインパクトの点で特に重要と思われる4つの課題について検討する。4つの課題とは、援助評価関心のプログラム・政策レベルへのシフト、国際的な教育援助協調の評価、「万人のための教育」の評価、そして途上国の自立的な発展に果たす評価の役割、である。

(1) 援助評価関心のプログラム・政策レベルへのシフト

前述したように、DAC 評価グループ・メンバーの現在の最大共通関心事は、援助の成果重視マネジメントである(DAC 2000)。援助機関の予算が縮小傾向にある中で、援助事業は全体としての有効性と効率性の証明を迫られており、必然的に個々の援助プロジェクトのレベルよりも高次の政策意思決定段階での評価が問題にされるようになってきている(DAC 1998)。既に見たように各援助機関の評価担当部局は、マネジメントあるいは議会の要請に対して実績モニタリングによる継続的な評価の実施で答えなければならない(USAID 2002; Effron 2002)⁽¹⁷⁾。このことは援助評価が一般的に政治性を増していく可能性を示唆している。教育援助、特に基礎教育に対する援助は人道的見地からどの援助国でも一定の好意的な評価を受けてきたが、教育援助評価には前述したような本来的な困難がつきまとうだけに、援助関心のプログラム・政策レベルへのシフトは教育援助評価にとって新たな挑戦となる可能性がある。

一方、成果重視の現場的な意味合いとして、過去に実施した事業から得られた知識や経験の系統的な蓄積を図るため、分野別のメタ分析による評価や国別援助政策のレビューのための評価の実施が必要とされている。世銀、米国、日本といった比較的援助規模の大きな機関の教育援助評価について見ると、USAID が既に多くの報告を刊行しているほか⁽¹⁸⁾、国際協力事業団も現在そのような調

査を実施中である。世銀は、OED を中心に実施事業の結果のデータベース化の一環として教育分野を扱ってきたが(例えば、Berk 2000)、2002年には上層マネージメントが教育分野の評価を優先すべきとの意思決定を行っている。ノールウェーのような援助小国の場合でも、個々の援助事業の規模が限られているだけに、全体的な効果に関心があり、より高次での援助評価を重要視する(長尾 2000; Nordic Consulting Group 2002)。各国で並行的に実施されつつあるこのような取組みが何らかの形で相互連携につながれば、教育援助の質的な飛躍につながる可能性も出てくる。

(2) 国際的な教育援助協調の評価

先進援助国の援助財政の逼迫は、援助協調を促進している。教育分野においては、1990年代に「万人のための教育」援助の大きな流れの中で、より積極的な意味でセクター・プログラムの形での協調が試されるようになった(King 1997)。セクター・アプローチには、通説的に事業対象がセクター(部門レベル)であること、セクター戦略が明確にされること、途上国側が主導すること、援助国側が資金のプールも含めて積極参加すること、途上国側との実施面での協調が図られること、そして可能な限り現地リソースを活用すること、の6要件がある。サハラ以南のアフリカ地域では、教育分野でも援助協調が始まっている。セクター・プログラムが有効に展開するには、途上国と援助国側の双方の情報ニーズを充足する協調評価体制を確立する必要があるほか、途上国側の評価能力を育成することが必須の要件とされる(DAC 1998)。問題は評価における協調よりも、セクター・プログラム事業そのものにおける協調が成立するかであるが⁽¹⁹⁾、評価プロセスを適切に組むことにより、事業における協調に対して促進的に働きかける可能性は検討すべき課題である。

(3) 「万人のための教育」の評価

Jomtien会議で始まった「万人のための教育(EFA)」の運動が「第2の10年」に入っ
て久しい。当初の10年間に途上国の教育開
発をモニターする体制が出来たことで、かつ
てHyneman(1999)が指摘したような教育統
計の概念的不備や統計収集方法の不一致等の
問題はかなり軽減され、教育開発・教育協力
について横並びで議論することが可能となっ
た。EFAのフォローアップについては、
UNESCOがモニタリングの役割を付与され、
少なくとも表面的にはEFAの継続的な
評価の態勢は整った。UNESCOが定期的に
発行するモニタリング・レポートは、各国、
特に先進援助国に対して「ダカール枠組み」で
合意した諸目標の達成に関するコミットメン
トの履行を促す役目を果たしている
(UNESCO 2002)。その間に、EFA実現のため
の実質的な国際協調プログラムとして世銀
の主導で開始されたFast Track Initiative
(FTI)が、2002年から動き出した。2000年
9月の国連ミレニアム・サミットで合意され
た教育目標(2015年までの初等教育の完全
普遍化と初中等教育における男女間格差の解
消)の達成に向けて、開発戦略の中に教育開
発を的確に位置づけていることと教育セク
ターの開発計画に関する援助国の同意が得ら
れていることの2条件を充たしている途上国
に優先的に資金を提供することにより、
EFA目標の早期実現を支援しようという計
画である。既に18の途上国が選ばれ、その
内の7カ国(ブルキナファッソ、ニジェール、
モーリタニア、ホンジュラス、ニカラグア、
ギニア、ガイアナ)がFTIプロセスに入っ
ている。これら7カ国で就学していない学
齢児童の合計は約390万人で、ダカール会
議で引用された1億1300万人の未就学
児童数からすると4%にも満たないが、と
にかくスタートが切れたのであるから、
これらに続くべき国や援助国の支援状
況も含めてこのプロセスをモニターして
いくことは大事である。18

カ国には入っていないが、全未就学児の半
数を抱える5カ国(インド、パキスタン、
バングラデッシュ、ナイジェリア、コン
ゴ民主共和国)についても特別扱いにつ
いて協議中であり、こちらも国際教育協
力の主要関心事項としてモニターされな
ければならない。前掲したUNESCOのEFA
レポートでも指摘しているが、ダカール
枠組みでの合意事項には、このような教
育の量的拡大に関する目標の他に、教
育の質の改善等の重要目標も含まれて
いる。EFAが単に教育開発の遅れた国
々の教育統計上の変化に終わらず、学
校レベルで生徒児童が受ける教育の質
の向上につながるかどうか評価の課題
として追跡する必要がある。EFAの量
的、質的なインパクトの本格的な評価
は、未だ端緒にもついている重要課題
である。

(4) 途上国の自立的な発展に果たす評価の役割

DACの援助評価5基準の中で最も中途
半端な扱いを受けているのが自立発展
性基準であろう。災害救援や難民救済
のような緊急援助は別として、通常
の援助事業では、少なくとも援助す
る側は、援助期間終了後に援助チ
ームが引き揚げた後でも、受入れ
途上国スタッフが事業を継続してい
けるようになる(あるいはする)こ
とを念頭に置いている。しかし、実
際には、途上国の側に事業の継続
に必要な予算措置を講じる余裕が
ない、事業継続に備えて育成した
人材が昇進、転職等で異動してし
まう、受入れ国側スタッフが「援
助漬け」になっていて、そもそも
自立発展を試みる意欲を持たない
等々様々な理由で、自立発展性を
確保するのは至難の業である。ま
た、FTIのように、EFA目標の早
期達成のために外部資金の導入を
ある意味で奨励するような事業
では、自立発展性基準は初めから
ある程度控えめに押さえてかか
らなければ、自己矛盾に落ちる危
険がある。もし、FTIで巨額の
援助を受けるとにより就学率目

標を達成したとして、「その後」はどのようなのであろうか。目標達成とともに支援を打ち切ると、それで就学率が再び低下する危険はないであろうか。もし日本が援助方針の中核に据えている「途上国の自助努力の支援」の原則に忠実に援助を実施するならば、事業評価における自立発展性基準の捉え方をより厳密に考える必要がある。前掲の世銀のICRでは「組織制度開発」と「自立発展性」の双方を目標基準として挙げている。しかし、教育事業のように時間軸が通常の事業よりも長期にわたる場合には、そのような評価慣行を考えるだけでは不十分で、より根本的な対応が必要かも知れない。例えば、事業の企画段階から自立発展性基準を重視する事業アプローチや組み立てを考えてはどうだろうか。自立発展性の発現に関する現実的な想定をもとに事業を組み立てるとすると、もしかすると現在であれば「3億円×5年＝15億円」で設計している事業を、「1.8億円×10年＝18億円」で作り直して実施する方が適当との解が出るかもしれない。少なくとも、最貧国の多くがそうであるように組織制度的開発の遅れが顕著な国々については、別扱いが順当と考えられる。援助における自立発展性基準の重視の意味合いが途上国の開発を通しての自立の獲得にあるとするなら、以上の議論は評価のあり方ではなく評価の役割と置き換えて考えるべきであろう⁽²⁰⁾。いずれにしても、これは長期的に重要な教育援助評価の新課題となるであろう。

6. 教育援助評価の今後の展開

以上4つの課題は、実は相互に複雑に関連している。それらの織り成す関係が教育援助評価の今後の展開を特徴付けることになる。援助評価をめぐる動きで最も顕著なのは、援助機関の成果重視マネジメント志向で(DAC 2000)、これはアカウンタビリティ目的の実績管理の標準化の形でほぼ全ての援助

機関で実現しつつある。アカウンタビリティ目的は、同時に評価の実施プロセスへの被援助国の利害関係者の参加を求めている。援助機関のマネジメントとしては、納税者の目を意識したスタンス的なところもあるが、現場の事業担当者や実際に評価に携わる専門家としては、個々の援助事業の置かれた社会的・文化的状況や利害関係者の状況に応じたきめの細かい対応を志向する。両者の間の意思の疎通が十分でないと、援助機関側の標準化圧力と現場の特殊文脈重視の姿勢の間で対立や葛藤が生じたとしても不思議はない。教育援助の場合には特に援助評価の現場の状況が複雑で、そのような関係がどう実現していくか不明瞭である。

これは援助評価の国際間連携や合同評価についても同様で、援助機関のマネジメントは原則として積極的支援の姿勢をみせているが、実際にはさまざまな問題がある。例えば、実施段階で関係する評価専門家が良心的に評価作業をしようとする、各参加援助機関の側でさまざまな対応が必要になる。中には標準化を志向するマネジメントと相容れない調整を迫られる可能性もある。個々の事業単位で、分散的に連携する範囲では問題は生じないであろうが、教育セクター・プログラムのようなかなりの規模の事業になると、目立つだけにマネジメントが合同評価の結果に対応できずに困るといった状況もありえる。また、連携・合同評価には確実に費用がかかるが、そのメリットはかならずしも一般市民にとって理解しやすくはない。その関連で費用対効果の注文が付く可能性もあろう。

EFAのフォローアップは、G8の検討議題になるなど、メディアに対しても目に付きやすい事業となっている。基礎教育ということでは人道的立場に立脚した議論がしやすいからであろう。その意味ではかなりの政治性を帯びた事業である。将来展開のシナリオとして念頭に置いておくべきことは、2015年が近づき初等教育の完全普遍化が実現できない可

能性である。そもそも世銀がFTIを設置するイニシアティブをとった理由のひとつが、現在の途上国の教育開発のペースでは、約80の国が2015年までにEFA目標を達成できないとの推測である。その場合には、モニタリング・評価の制度自体が役に立たなかったとの批判を受けることになるかもしれない。特にEFAの教育現場での質的なインパクトを把握していないと、二重に信頼性を失う機会を提供してしまうかもしれない。そこまで極端に悲観的な展開を想定しないとしても、EFAプロセスでの援助依存に対する疑問や批判は当然出てこよう。それは途上国の教育開発に対する援助が、そもそも人材開発の支援を通じて途上国の自立的発展に貢献するとの暗黙の理解があるからである。

EFA・FTIに対する批判を待つまでもなく、途上国の自立的な発展とそのための教育の役割は途上国、先進国を挙げての重要課題となる可能性がある。なぜなら教育開発の最も遅れているサハラ以南アフリカ諸国が、2001年に「アフリカ開発のための新パートナーシップ(New Partnership for Africa's Development /NEPAD)」を宣言し、アフリカ諸国が相互に連携して自立的・持続的発展可能路線を模索していくことと、教育・人材育成を通じる人間開発を中心的課題と位置づけることについて明確な意思表示を行っているからである。この新しい動きに、第3節でレビューした評価慣行のアフリカへの普及を重ね合わせて考えると、教育評価における将来的な展開として、途上国側での自立発展性重視の評価の積極的活用が考えられる。援助する側としてもそれに対応すべく個別事業の評価に関する自立発展性基準の適用について真剣に検討すべきであろう。

謝辞：本稿の執筆に際し、榎本伸悦氏(広島大学大学院国際協力研究科)に教育援助評価資料の収集と予備的な分析について支援をいただいた。ここに記して謝意を表したい。但

し、言うまでもなく文中の誤りは全て筆者に帰するものである。

注

- (1) DAC評価5基準の設定の背景および用語の定義については、DAC Working Party on Aid Evaluation (1986, 1991, 2000)を参照されたい。
- (2) 本稿では、Popham(1993)に従って、教育評価を「教育事業・活動の質の系統的な検証」と定義する。
- (3) 以下の教育評価の歴史的展開に関する記述は、主にWorthen and Sanders (1987, pp. 11-20)およびMadaus and Stufflebeam (2000, pp. 3-18)に依拠している。
- (4) 教育評価の及ぼした最大の影響は、大学院レベルでの評価教育普及の母体となったことであろう。Altschuld et. al. (1994)の1993年の調査に依ると、事業評価を専門とする修士・博士課程プログラムを有する38校の米国の大学院のうち、25校(約2/3)は教育学・教育心理学・心理学系プログラムであった。
- (5) 日本における教育評価の歴史については、梶田(1999)および井上(2000)を参照されたい。
- (6) 本論集でも齋藤・黒田(2000)が研究成果の一部を紹介している。
- (7) アフリカ評価学会の評価基準ドラフトについては、同協会のホームページ(<http://www.geocities.com/afreval/documents/aeg.htm>)を参照されたい。
- (8) IOCEについては、同機関のホームページ(<http://home.wmis.net/russon/ioce/>)を参照されたい。
- (9) 広島大学教育開発国際協力研究センターでは、現在各国の援助機関が教育援助事業をどのように評価しているか、またその過程でどのような問題に遭遇し、それ

- とどう対処しているかについて、横断的な研究調査を実施中である。
- (¹⁰)以下の世界銀行の評価慣行に関する記述は、2003年5月19～20日に筆者が世界銀行を訪問して行った聞き取り調査(含ICRガイドライン等の部内資料の閲覧)に基づく。
- (¹¹)例えば、「かなり満足」な成果は、「事業が全ての主要関連目標を達成したか、超過達成し、かなりの開発効果を既に実現したか、かなり実現する見込みがあり、とりたてて問題のない」場合。「不満足」は、事業がほとんどの主要関連目標を達成できず、目立った開発効果も実現しておらず、これからも実現の見込みがない上に、重大な問題を抱えている」場合である。
- (¹²)「初中等教育投資に対する収益は10～15年後にようやく実現する不確定要素の大きい変数であり、事前投資予測値を変更する統計が得られないので、あえて計算する意義が見出せない」との判断である。
- (¹³)「満足」の側には「わずかに満足」が、「不満足」の側には「わずかに不満足」が含まれる。また「満足」、「不満足」の付く3段階間での修正は考慮に入れていない。前述したようにESが全てのICRを対象とするのに対して、PPARの件数は限られている(1998年度が約60件、2000年度が23件)ので、3次評価はその範囲での意味しか持たない。
- (¹⁴)成果の判定だけでなく、組織制度的開発、自立発展性、世銀と融資先のパフォーマンスの判定についても、同様のことが言えることで、PPAR実施の必要性は高いとも指摘している(Effron 2002, pp.38-39)。
- (¹⁵)以下のDFIDに関する記述は、同機関のホームページに掲載してある"DFID's Evaluation Policy"(July 2002)、「Draft Evaluation Guidelines」、及びダウンロード可能な教育援助事業評価報告書(DFID 1998)に依拠している。
- (¹⁶)以下のUSAIDに関する記述は、2003年5月19～20日に筆者が同機関を訪問して行った聞き取り調査およびUSAID(2003)に基づく。
- (¹⁷)国際協力事業団でも、2003年10月の独立行政法人化に向けて、実績評価管理システムの導入を準備中である。
- (¹⁸)USAIDのホームページのDevelopment Exchange Clearinghouseを参照されたい。
- (¹⁹)世銀での聞き取り調査(上記注10を参照)によると、ザンビアの基礎教育セクター・プロジェクト(1999～2002年)を支援している13の援助国が2002年に共同で実施した評価では、援助国側と途上国側の関係はかなり改善し、協調アプローチの面では前進が見られたが、援助国側が財布のひもを握る構造は変わらず、また援助国と途上国政府の関係が出資方法で変わる等の事情で、協調の進展にも問題が残るようである。
- (²⁰)評価を個別事業の展開のツールとして活用する試みについては、長尾(2001)、牟田(2003)を参照されたい。

参考文献

- Altschuld, James. et. al. (1994) "The 1994 Directory of Evaluation Training Programs" in Altschuld, J. and Engle, M. (Ed) *The Preparation of Professional Evaluators: Issues, Perspective, and Programs*, New Directions for Program Evaluation, No. 62, San Francisco: Jossey-Bass Publishers, pp. 71-94.
- Bamberger, Michael (2000) "The Evaluation of International Development Programs: A View from the Front" *American Journal of Evaluation*, Vol. 21, No. 1, pp. 95-102.
- Berk, David (2002) *Tertiary Education:*

- Lessons from a Decade of Lending, FY1990-2000*, World Bank, Washington, D.C.
- Bhola, H. S. (2003) “Social and Cultural Contexts of Educational Evaluation: a Global Perspective” in Kellaghan, Thomas and Stufflebeam, Daniel L (Ed.) *International Handbook of Educational Evaluation*, Dordrecht: Kluwer Academic Publishers, pp. 397-415.
- DAC (1986) *Glossary of Terms Used in Evaluation*, Working Party on Aid Evaluation, Paris: OECD.
- DAC (1991) *The DAC Principles for the Evaluation of Development Assistance*, Working Party on Aid Evaluation, Paris: OECD.
- DAC (1998) *Review of the DAC Principles for Evaluation of Development Assistance*, Working Party on Aid Evaluation, Paris: OECD.
- DAC (2000) *Results Based Management in the Development Co-operation Agencies: A Review of Experience, Executive Summary*, Working Party on Aid Evaluation, Paris: OECD.
- DAC (2002) *Glossary of Key Terms in Evaluation and Results Based Management*, Working Party on Aid Evaluation, Working Party on Aid Evaluation, Paris: OECD. (和訳：経済協力開発機構開発援助委員会、援助評価作業部会，2002 .『評価と援助の有効性 - 評価および結果重視マネジメントにおける基本用語集』)
- DFID (1998) *Strengthening Primary Education in Kenya, An Evaluation of the Strengthening Primary Education (SPRED) Project, Kenya 1991-1996*, Evaluation Study (EV627), London.
- Effron, Laurie (2002) *2002 Annual Report on Operations Evaluation*, The World Bank, Washington, D.C.
- Garaway, Gila (2003) “Evaluating Educational Programs and Projects in the Third World” in Kellaghan, Thomas and Stufflebeam, Daniel L (Ed.) *International Handbook of Educational Evaluation*, Dordrecht: Kluwer Academic Publishers, pp. 701-720.
- Hyneman, Stephen (1999) “The sad story of UNESCO's educational statistics” *International Journal of Educational Development*, 19, pp. 65-74.
- 井上正明編，2001 『教育評価読本』教育開発研究所 .
- International Consultative Forum on Education for All (2000) *Education for All 2000 Assessment: Global Synthesis*, Paris, UNESCO.
- 外務省，2003 『経済協力評価報告書 2002』ホームページ
(www.mofa.go.jp/mofaj/gaikou/shiryou)
- 梶田叡一，1999 『教育評価』有斐閣 .
- Kellaghan, Thomas and Stufflebeam, Daniel L (Ed.) (2003) *International Handbook of Educational Evaluation*, Dordrecht: Kluwer Academic Publishers.
- Kellaghan, Thomas and Stufflebeam, Daniel L. (2003) “Introduction” in Kellaghan, Thomas and Stufflebeam, Daniel L (Ed.) *International Handbook of Educational Evaluation*, Dordrecht: Kluwer Academic Publishers, pp. 1-6.
- King, Kenneth, 1998. “New Challenges to International Development Cooperation in Education” 『国際教育協力論集』 1 巻 1 号，3-16 頁
- 国際協力銀行，2002年『円借款案件事後評価報告書 2002』東京 .
- 国際協力事業団，2001 『JICA 事業評価ガイドライン』東京 .

- Madaus, George F. and Stufflebeam, Daniel L. (2000) “Program Evaluation: A Historical Overview” in Stufflebeam, Daniel L. et. al. (Eds) *Evaluation Models: Viewpoints on Educational and Human Services Evaluation*, 2nd Edition, Boston: Kluwer Academic Publishers, pp.3-18.
- 三好皓一, 2003. 「特集にあたって(政府開発援助(ODA)における評価の新たな潮流)」 『日本評価研究』 3巻1号.
- 牟田博光, 1998. 「教育ODAの経験と課題」 『国際教育協力論集』 1巻1号、17 - 30頁.
- 牟田博光, 2003. 「構造的評価に基づく総合的国際協力の試み」 『日本評価研究』 3巻1号.
- 長尾真文, 1999. 「教育援助に関する研究課題」 『国際教育協力論集』 2巻2号、143-154頁.
- 長尾真文, 2001. 「援助における評価の目的と活用方法: 南アフリカ理科教育支援事業による例示」 『国際教育協力論集』 4巻1号、89-100頁.
- Nagao, Masafumi (2000) “The role of evaluation in Norway's development assistance”, mimeo.
- NEPAD (2001) *New Partnership for African Development*, (URL: <http://www.un.org/esa/africa/NEPADenglish.pdf>.)
- Nordic Consulting Group (2002) *Norwegian Support to the Education Sector: Overview of policies and trends, 1988-1998*, Oslo.
- Omolewa, Michael and Kellaghan, Thomas (2003) “Educational Evaluation in Africa” in Kellaghan, Thomas and Stufflebeam, Daniel L (Ed.) *International Handbook of Educational Evaluation*, Dordrecht: Kluwer Academic Publishers, pp. 465-481.
- Picciotto, Robert (1997) “Evaluation in the World Bank: Antecedents, Instruments, and Concepts” in Chelimsky, Elenor and Shadish, William (Ed.) *Evaluation for the 21st Century: A Handbook*, Thousand Oaks: Sage Publication, pp. 201-213.
- Postlethwaite, T. Neville (1999) *International Studies of Educational Achievement: Methodological Issues*, Hong Kong: Comparative Education Research Centre, Hong Kong University.
- Psacharopoulos, George (1995) “Using evaluation indicators to track the performance of education programs” in World Bank, *Evaluation and Development*, Proceedings of the 1994 World Bank Conference, Washington, D. C.
- Reimers, Fernando (2003) “The Social Context of Educational Evaluation in Latin America” in Kellaghan, Thomas and Stufflebeam, Daniel L (Ed.) *International Handbook of Educational Evaluation*, Dordrecht: Kluwer Academic Publishers, pp. 441-463).
- 齋藤みを子・黒田一雄, 2000年. 「アフリカ7カ国における初等教育就学児童の読解力の男女間格差に関する統計的考察 - 教育の質調査のための南アフリカ諸国連合(SACMEQ)の調査結果から - 」 『国際教育協力論集』 3巻1号、25-39頁
- Svensson, Kristina (1997) “The Analysis and Evaluation of Foreign Aid” in Chelimsky, Eleanor and Shadishi, William (Ed.) *Evaluation for the 21st Century: A Handbook*, Thousand Oaks: Sage Publication, pp. 272-283.
- UNESCO (2002) *EFA Global Monitoring Report 2002*, Paris.
- USAID (1996a) *Performance Monitoring and Evaluation: Selecting Performance Indicators*, Center for Development Information and Evaluation, No. 6, Washington, DC.
- USAID (1996b) *Performance Monitoring*

- and Evaluation: Preparing a Performance Monitoring Plan*, Center for Development Information and Evaluation, No. 7, Washington, DC.
- USAID (2003) *USAID Performance and Accounting Report FY 2002*, Washington, DC.
- Wiesner, Eduardo (1997) "Evaluation, Markets, and Institutions in the Reform Agenda of Developing Countries" in Chelimsky, Eleanor and Shadishi, William (Ed.) *Evaluation for the 21st Century: A Handbook*, Thousand Oaks: Sage Publication, pp. 189-200.
- World Bank (2002) *Project Performance Assessment Report: Brazil (Innovation in Basic Education Project, etc.)*, Report No. 24433, Washington, D.C.
- Worthen, Blaine R. and Sanders, James R. (1987) *Educational Evaluation: Alternative Approaches and Practical Guidelines*, New York: Longman.
- 横関裕見子, 1999 「サハラ以南アフリカ地域の教育と教育セクター・プログラム」 『国際教育協力論集』2巻1号、101-112頁

“Present concerns and issues in evaluation of educational cooperation”

Masafumi Nagao

The global attention to 'Education for All' during the 1990s has also highlighted the importance of monitoring and evaluation in the area of educational development and cooperation. In this paper the author examines if this recognition has given rise to the establishment of its practice, through a review of principal concerns and issues in the relevant literature. The review starts with a look at recent developments surrounding aid evaluation, followed by consideration of development and international diffusion of the practice of educational evaluation. The paper then examines the internal evaluation system established by the World Bank for assessing the outcome of its projects, and then reviews the corresponding practice in USAID, DFID and JICA. It also discusses 'new' issues, including the shift in evaluation focus toward programs and policies (away from projects), joint evaluation practice, evaluation of EFA, and evaluation of sustainability. Although the spread of results-based management practice and the development of EFA follow-up are drawing the evaluation concern closer to the head offices of aid organizations, field-level concerns are equally important, especially since developing countries are becoming interested in evaluation practice as a way of assuming greater initiative in development.