

公民館の評価をめぐる理論的論点

佐々木 保孝

はじめに

本稿は、近年、社会教育関係者の関心を集めている「評価」を論じる際の視角に関する一考察である。念頭には、社会教育における中心的施設である公民館を置いているが、理論的な検討を中心におこない、ひろく社会教育活動の全般に敷衍できる要素について考察していきたい。

公民館を含む公共施設の経営について評価する場合、例えば参加者数や施設の稼働率といった数量化された指標でもって運営状況の評価する傾向が強くなっている。これは、行政全体の動向として、評価活動が「新公共管理（New Public Management）」と総称される行政改革の思潮のもとで進められたことによる。具体的には、政策の効果や効率性などに注目し、それらの側面を数量的に把握するための「業績測定（performance measurement）」が評価手法の主流となったことの影響が大きかった。

このことは、公民館の現場において、従来行われてきたような事業のふりかえりや住民との対話という“評価”のあり方と大きなギャップをもつものとして受け取られた。だからこそ、千葉県や福生市をはじめ全国各地で模索を続ける公民館においては、自分たちの手で事業評価シートの作成が試みられた¹。また、公民館の評価を正面から論じるものとして、2000年代の中ごろから後半にかけては、『月刊公民館』や『月刊社会教育』誌上で何度か特集が組まれ²、関連する論文が発表されるなど³、公民館評価に関する研究の機運も高まりをみせている。とはいえ、公民館の評価をめぐるのは、まだ研究の数自体が少ないうえに、その大半は、個別事業の評価方法を論じることに主眼がおかれている。確かに事業評価法の開発は具体的に求められるところであるが、上野が指摘するように⁴、各地で多様な様態をもつ公民館であってみれば、各々で開発された手法がそのまま一般化できるとは限らない。目下の状況は、公民館評価のための理論開発もまた同様に急務なのである。

そこで注目されるのが、主としてアメリカ合衆国で発展を遂げた「プログラム評価 (program evaluation)」論である。その歴史をひも解くと⁵、1930年代のニューディール政策下であるとか、1960年代のジョンソン政権下など、公共政策に巨費が投じられた時代に政策の効果について社会的な関心が高まった際、理論や手法に発展がみられた。特に、ジョンソンが「貧困との戦い(War on Poverty)」のスローガンを掲げて、社会問題解決のための諸プログラムを矢継ぎ早に実施ことで、評価の問題に注目が向けられた。関心の中心は、特定の社会問題の改善に対して打ち出した諸々の個別事業が総体として成し遂げた成果とそれに至る過程にあったという。筆者は共著の別稿においてその理論の基礎的なポイントを概観したが⁶、本稿ではその補足的続編として、そこから導かれる公民館評価の論点についてさらに考察を進めてみたい。

1. プログラム評価論の特徴

プログラム評価論の分析にあたって、タームを確認しておこう。本邦の社会教育において「プログラム」という場合、公民館等の施設において開催されている個々の講座を指すことが多い。行政評価でいうところのいわゆる個別事業にあたるもので、現場レベルで個別事業評価が求められる事情とも相俟って、講座の学習目標、学習内容、学習方法、講師、受講の満足度といったものに事業予算等を加味した評価の指標(評価シート)の開発に力が注がれている現状がある。

しかし、公民館評価にプログラム評価論の原理や技法を応用しようとする際には、「プログラム」を個別事業に限定しない方が活用しやすい。いみじくも、成人教育学者のノックス(Knox, A. B.)は、ニーズの把握に始まり、目標の設定、予算、スタッフ、参加、提供物(program)、施設・設備から、学習の成果と効用にいたる過程全体を「教育プログラム」と呼び、総体として評価を行うことを提唱した⁷。これまで本邦の社会教育が、政策、行政体系、業務、施設経営、そして学習成果と、全体の一部を切り取り、脈絡を欠いたまま評価を行うのを常態としてきたことを思うと、「プログラム」＝個別事業とイメー

ジを限定してしまうことは、プログラム評価論の理解をかえって難しくする。よって、ここでは、当該の公民館がおこなう諸活動 —主催事業、施設におけるグループの学習活動、学習相談、学習団体・学習者の組織化等— の総体をひとまず「プログラム」と捉えることとして考察に入ろう。

プログラム評価論では、先述したように、プログラムの成果と過程に主たる関心が向けられる。その手法を大づかみにいえば、まず、検証される「成果」とは何か、それはどのような筋道で達成されるのかということ进行を想定し、実際の状況を検証するというものである。言いかえれば、成果とその達成の道筋についての想定作業無しには、評価自体のやりようがないこととなる。

この筋道のことを「プログラム・セオリー (program theory)」とよぶ。「セオリー」とは日本語の「理論」という語感よりも、英英辞典のなかに「なぜ起ったのか、あるいはなぜ存在しているのかについて説明する意図をもった一連の公式な考えのまとまり」とあって、こちらの方がここでの意味合いに近い。すなわち、「プログラム・セオリー」とは、ある原因がある結果を引き起こし、それが原因となって次の結果が起こるという原因と結果が連鎖上につながる仮定をプログラムの中に見出したものとなる⁸。

プログラム・セオリーを最も単純にモデル化すると、基本形は、【〈活動をする〉→〈成果を得る〉】となる。たとえば、「健康の改善」という成果を設定すると、【〈適度な運動〉→〈健康の改善〉】といった関係になる。これを、ケロッグ財団のプロジェクトにおいて長年評価活動に従事したウエスタン・ミシガン大学のノールトン (Knowlton, L. W.) とフィリップス (Phillips, C. C.) は、「変化のセオリー・モデル (Theory of Change Model)」とよんだ⁹。〈活動〉の部分は当然ひとつとは限らないから、上記の例で言えば、【〈適度な運動〉／〈栄養の改善〉／〈ストレスの軽減〉／〈体力の増強と保持〉】の4つの戦略的な活動を原因として【〈健康の改善〉】という結果を得るというモデルになりうる。変化のセオリー・モデルの〈活動〉と〈成果〉に当たる部分をさらに構造化すれば、より詳細なプログラムのセオリーを描き出せるが、この点については後述する。

公民館では、個別事業ではなく活動の全体を対象に「プログラム・セオリー」を記述すること自体、これまでほとんど試みられなかったことであろう。それだけでも示唆は大きいと思われるが、ここで最も押さえておきたいことは、セオリーを考える手順として、プログラムの成果を想定するところから始めている点である。

「成果の想定」とは、教育関係者にはなじみやすい考え方もかもしれない。なぜなら、学習計画の策定において到達目標（goal）を設定する際、それは成果の想定という側面をもつからである。しかしながら、公民館評価の現状に足りない部分を考えてとき、なじんでいるからこそ気づきにくいもっと別のところに、「成果の想定」という発想の利点を見出す必要があるのではないか。

それは、ふだん、施設の管理から講座の運営、学習支援を通じた学習者と職員との交流など、様々な活動が複層的に営まれている公民館活動の日常に、ある特定の成果をひとつの視点として設定し、そこからみると当該の公民館がどのように機能しているのかが可視化できるといったイメージである。つまり、「健康の改善」という成果を想定したとき、例えば、当該の公民館には関連講座が幾つあってどんな学習が展開されているのか、自主グループの中でそういった取り組みはあるのか、職員に学習者はその問題に関連するどんなコミュニケーションをとっているのか、そして、そうした諸活動がどのように関連して最終的に成果の達成に向けて作用しているのか。そういった姿が浮かび上がってくるのである。

当然、別の視点を用いれば異なる姿が見えることになる。プログラム評価論の著名なハンドブックを作成しているロッシ（Rossi, P. H.）らは、「プログラム評価の最も難しい側面のひとつは『すべてに適合する（one size fit all）』のアプローチが存在しないことである」とのべている¹⁰。すなわち、プログラム評価は、想定される成果によって異なる中身が構成されるものであり、ロッシらの表現をつかえば、本来は個別に「あつらえ」られ（tailored）なければならないのである¹¹。

こうした点をプログラム評価論の第一の特徴とすれば、第二には、評価結果

の活用を意図したプログラムの利害関係者（stakeholders）の存在が注目される場所である。ロッシらは利害関係者を「あるプログラムがどの程度機能しているかに対して重大な関心をもっている個人、集団、または組織のこと」¹²と定義し、具体的に考慮する必要がある者として、「政策立案者と意思決定者」、「プログラムスポンサー」、「評価スポンサー」、「標的となる参加者」、「プログラム運営者」、「プログラムスタッフ」、「プログラムの競合者」、「社会文脈上の利害関係者」、「評価と研究の学界」を挙げている¹³。一見ただけで多様な人たちが想定されていることが分かるが、評価を行う場合は、プログラムに関心を寄せる利害関係者を同定し、できるだけ早くから関わることによって要望を的確に捉え、それに応えうる評価をみつらえるべきであるとする。

利害関係者に応じて何のためにどのような評価結果を欲しているのかが異なってくるから、その点は評価担当者と利害関係者がよく話し合い、焦点を絞っていかなければならない。これを「評価クエスチョン」の設定とよぶが、この問題について、ロッシらが分かりやすい例を提示しているので、やや長い引用しておこう。

「たとえば、地域保健プログラムの運営者は最初、『私たちの教育およびアウトリーチサービス活動は、人々にAIDSのリスクについて伝えることに成功しているだろうか？』と尋ねてくるかもしれない。しかし実際は、そのサービスとは、市民クラブの会合や健康祭りにプログラムスタッフがたまに出かけて行って話をするくらいのもかもしれない。このようなかなり小さな活動レベルでは、幅広い人たちがAIDSに関して多くの情報を得ることを期待するほうが非現実的であろう。もし仮に、このサービスに関して評価クエスチョンを考えるとすれば、それはたとえば『私たちの教育およびアウトリーチサービス活動は、そこに参加する人たちのAIDS問題に対する意識を高めるだろうか』とか『その参加者は、他の人たちのAIDS問題に対する意識に影響を与えそうな地域のリーダーたちを代表しているだろうか』といったもののほうがよいだろう。（中略）その教育およびアウトリーチサービス活動が控えめであることは、明らかに『人々にAIDSのリスクに関して伝える』ところまで至っておらず、

これを実際のアウトカムと評価してもほとんど意味がないであろう。評価者と利害関係者は、プログラム実績に関連してくるプログラム構成要素、活動、人的配置を精査して同定し、それらの特徴に対して妥当な評価クエスチョンを形成すべきである。」¹⁴

評価の枠組みづくりにあたって、評価クエスチョンの回答可能性を十分に検討することを、プログラム評価論の三点目の特徴として指摘しておこう。評価クエスチョンはできるだけ曖昧な表現を避け、観察可能な指標で構成されることがポイントとなる。また、評価に充てる期間、評価者の資源や能力等によっても調整が必要な面があることにも留意しておこう。

2. プログラムの構造化とその活用

プログラムが成果をあげるためには、成果を想定した計画を立てられることがまず重要になる。しかし、いくら立派な計画をたてても、それが実施されなければ意図した成果にはたどりつけないであろう。むろん、思いがけない成果が出ることはあるけれども、プログラム運営の質を高めるノウハウを蓄積しようとするれば、実施の過程でどのような動きがあつて、意図通りに、あるいは意図しない形でプログラムが進行したのかという点を検証することが重要となる。つまり、プログラムの機能や作用に関して情報を収集し、改善のためのポイントを判断する「プロセス評価」が求められるわけである。

先述したように、評価にあたっては、その前提条件として、成果とそれに関連する道筋を「原因－結果」（あるいは、「もし－それなら：if-then」、「どのように－なぜ：how-why」等でも可能）の関連で示した「プログラム・セオリー」を策定することが求められる。プロセス評価のポイントは、プログラム・セオリーと実際のプログラムの作用のズレを検討することにあると言える。

そこで、プログラム・セオリーの策定について、ノールトンらの説明をもう少し追ってみよう。彼女たちは、「変化のセオリー・モデル」でまず大枠を組み立ててから、その構成要素を細分化する形でセオリーを描くことを提唱している。その関係を示したものが図1である。

図1の上部が「変化のセオリー・モデル」で、タームに厳密に従えば、「おこなうこと (Do)」の部分で「戦略 (Strategies)」、「得るもの (Get)」の部分で「結果 (results)」としている。前者から3つの要素が、後者から4つの要素が細分化され、それらをつなげたものが図の下部で、これは「プログラム・ロジック・モデル (Program Logic Model)」といわれる。プログラム・ロジック・モデルそのものは、ノールトンらのオリジナルというわけではなく、広く一般に言及されているものである。

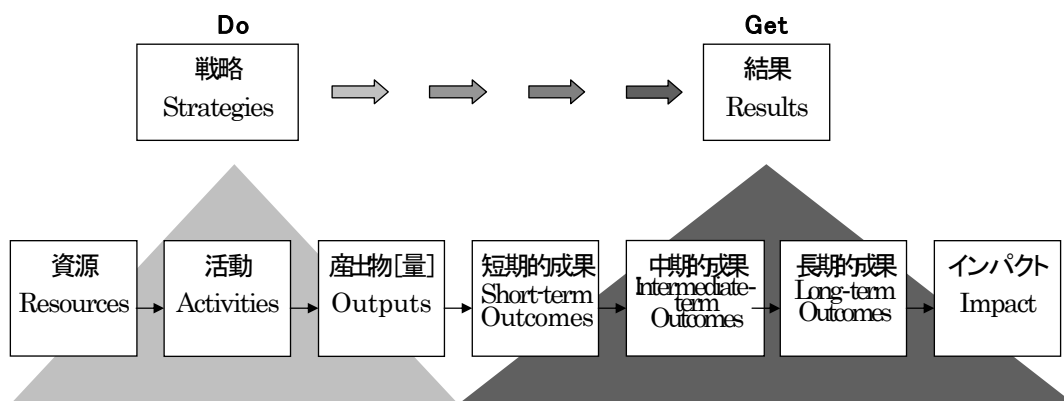


図1 〈変化のセオリー・モデル〉と〈プログラム・ロジック・モデル〉の関係¹⁵

プログラム・ロジック・モデルの前半部（戦略より降りてきた部分）の構成要素は、「資源 (Resources)」、「活動 (Activities)」、「産出物[量] (Outputs)」である。「資源」は活動をおこなうのに必要な人的、物的、財的、組織的、あるいはシステムに関するもので、〈投入 (inputs)〉と表記されることもある。「活動」には当該のプログラムで意図した成果をだすための種々の個別事業が含まれ、総体としては〈介入 (intervention)〉と呼ばれることもある。「産出物[量]」は文字通り活動が何を生み出したかを記述したもので、講座の受講者数や補助の受給額の推移など、量的にはかることのできるものが相当することが多い。なお、原語の output は次の outcome (成果) と区別して〈結果〉と

訳している和書も多いが、ここでは、results の訳語と区別し、かつ意味をより正確に表す意図から「産出物[量]」としておこう。

図の後半部分（「結果」より降りてきた部分）は、大きくは「成果（outcomes）」と「インパクト（impact）」から成っていて、成果については、短期・中期・長期のスパンに分かれる。その境はプログラムの中身によって相対的なものではあるが、目安として、短期で1～3年、中期で4～6年、長期で7～10年程度となる。多くのケースでは、プログラムに参加した個人や組織が知識や技術、意識や態度といった面で辿った変化を記述する。「インパクト」は、意図した変化について、それ以上は分析できない最終の状態で、個人レベルを超えて、組織やコミュニティについて言及することもある。

あるプログラム（活動の総体）について評価をしてみようと思えば、上記の諸要素に具体的状況をはめ込んで、左から右へ矢印でつなぐことができるかを確認すればよい。例えば、インパクトに「地域の健康意識の高まり」をあげたとしよう。すると、成果の部分に健康づくりに関する知識を得て、そこから近隣の人々と問題を考えていきたいという意識が芽生え、実際に健康づくりのイベントを企画する人が増えるというような、短期・中期・長期の成果を想定することができる。このようになるためには、「活動」として実際に健康づくりについて学べる講座が行われていなければならないし、それを実施しうる資金や講師、コーディネーター、受講者などが「資源」として調達されなければならない。講座が開かれると、成果からインパクトにつなげるためには、例えば当該地域人口の10%には最低でも受講してほしいといったように「産出量」を数字の根拠とともに想定しておく。このように、すべての要素が満たされれば、評価を行ってもある程度の結果が得られる可能性が高いと言えるが、例えば、活動が想定している成果を得るのにはかなり小規模であるとか、資源として講師が確保できないなど、図中の矢印が途中で途切れてしまうことはありうる。そうなる、その枠組みでは成果やインパクトを評価することはできないから、改めて別のアイデアを構想していくことになる。

以上のような作業を経て妥当なプログラム・セオリーを策定しておけば、実

際にプログラムを実行した段階で、どこに課題や利点があったのかがはっきりする。こうしてプロセスのブラックボックス化を防ぐことで、評価の目的であるプログラムの改善、あるいは利害関係者に対する説明責任を果たすことができる。さらには、プログラムの成果を検討するのに必要なプログラム実施中の状況をモニタリングする、という意味もある。

さて、ここまでは、プログラムの構造に関する側面を中心に論じてきたが、やや観点を変えて、プログラムの主たる対象者（target）の問題を検討してみよう。事業の対象者はあらゆる対人サービスを志向するプログラムには必ず存在するが、特に公民館は地域に根差していることでその意義を発揮する機関であるから、当該地域にどんなニーズがどの程度あるのかを把握するは極めて重要である。例えば、子育て支援であれば、対象地域の中にどれぐらい子育て世帯があるのか、子育てに関わる仕事をしたり団体活動をしている人はどれぐらいいるのかといったことは、プログラムを立ち上げるうえでの基本情報となる。社会教育の領域で、当該地域の学習ニーズを把握するための調査手法などが一定の蓄積をみている所以はここにある。

これを、プログラム評価として考えたとき、プログラムが適切なターゲット集団に行き届いているかどうかを査定するのも、評価の主要な役割のひとつである。いわゆる、プログラムの適切な供給を査定するカバレッジ（coverage：到達範囲）の分析と呼ばれる機能がそれである。カバレッジをめぐる問題は、プログラムが提供するサービスが需要よりも少ない（つまり、サービス供給量が足りない）過小カバレッジと、反対に需要よりも供給過多である過大カバレッジがある。

過小カバレッジは、需要予測が難しく、対象者数の把握に失敗した際におこるケースが多い。ただし、たいていのプログラムでは正確な需要予測をおこなうことが難しいことをふまえれば、むしろ、プロセス評価によって実態をモニタリングしながら、一定の条件で優先順位をつける、あるいはプログラムの期間を延ばすなどして、対象者数をプログラムが運営可能な状態にまで調整する必要がある。

過大カバリッジによる供給過多については、予算をはじめとした資源の無駄につながるほか、質的な面でも問題がおこることがある。例えば、自然体験の乏しい子どもたちを対象に組んだプログラムに余裕があり、自然に慣れ親しんだ子どもたちも数多く参加したとすれば、両者のギャップを埋めようというプログラムの意図は達成されないこととなる。評価によってその点が明らかになれば、対策としては、供給量を調整したり、あるいは、ねらいや内容に工夫を加えて（上記の例えで言えば、自然体験が豊富な子どもにそうでない子どもを手助けする役割を与え、協調的な態度を通じて両者のギャップを埋めるなど）プログラムを洗練させていくという方向もあろう。

いずれにしても、カバレッジの程度は、プログラムの記録を分析したり、新たな調査を行うことで把握していくことになり、社会科学の手法を精緻に活用していく必要がある。

3. 「アウトカム」の概念と評価のための課題

我が国におけるプログラム評価論の研究書では、成果の評価に関しては「アウトカム評価」と記載されるものが多いようである。「成果」という用語にすでに手垢がついている感覚であるとか、あるいは、先述したロジック・モデルにおいて、**output** と **outcome** の違いを強調するためなど、なんらかの意図はあるようだ。しかし何より、概念に馴染みがうすいことが、「アウトカム」という表記を採用する要因ではなかろうか。本節でも「アウトカム」と記すが、そのコンセプトがもつ意味について検討を加えてみたい。

ロッシらは、ハンドブックのアウトカムを扱った章で、まず、「アウトカムとは、あるプログラムが変化をもたらすことが期待されている標的集団や社会状況の状態である」と述べ、「対象とした人たちや社会状況に関する観察された特性であり、これはプログラムの特性ではない」と強調する¹⁶。これを噛み砕いて考えれば、例えば、5回構成の「万葉集講座」にすべて出席した事実はアウトカムではなくアウトプット（プログラムが産出した結果）である。そして、参加者が万葉集に関する知識を身に付けた状態がアウトカムということ

にある。確かに、万葉集の知識が身に付いた状態というのは当該の学習者が獲得した特性であって、万葉集講座の特性ではない。

さらに、ロッシらはこう続ける。「別の言い方をすれば、アウトカムとは、原則としてプログラムを受けなかった人たちや状況においても常に観察される特性に関することである。」¹⁷ 同じ例で考えれば、確かに万葉集の知識の有無はプログラムを受けていない人にも観察できる特性である。しかし、アウトカムを「成果」と考えれば、プログラムを受けていないにもかかわらずアウトカムがあるというのはやや奇異にも聞こえる。ここで彼らが言うアウトカムとは「状態・状況」といったニュアンスである。そのようにアウトカムの概念を設定したのは、次の図2を説明するためであった。

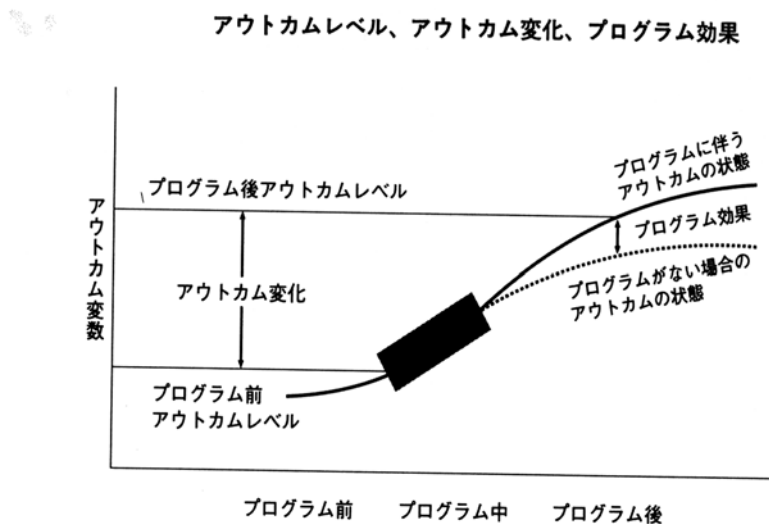


図2 アウトカムレベル、アウトカム変化、プログラム効果¹⁸

図の縦軸にある「アウトカム変数」とは、プログラムの活動によって影響を受けるプログラムの標的集団に関する測定可能な特性や状況（例えば、知識量

など)である。横軸は時間の経過を示す。アウトカム(標的集団の状況)をプログラムの実施前後で比較すると、曲線部で表されているように、その変化が観察できる。曲線のうち、点線はプログラムがない場合のアウトカムの状態で、例えば、万葉集に関する特集のテレビ番組を見ていれば、講座を受けていなくても知識量は増えているかもしれない。よって、実線で示されたプログラムに伴うアウトカムの状態から点線部を引いた状態がプログラムの純効果(実質効果)ということになる。本稿のはじめに、そもそも「プログラム評価論」が発達した背景にプログラムの効果を把握したいという点があったと述べたが、これを理論的に構造化したのが図2であるとも言える。つまり、慣習的に「成果」という意味で用いるアウトカムは、純効果を指している。

前節の図1で示したロジック・モデルにおいて、アウトカムを短期・中期・長期に分けていたが、これは、プログラムの純効果がどの段階でどのようにあらわれているのかを設定していると理解したほうがよい。長期的アウトカムからインパクトにかけてはプログラムが最終的に目指す成果にあたるが¹⁹、当然、プログラム終了時から遠くなるにつれてプログラム以外の影響要因が増えるわけであるから、純効果は測定しにくくなる。よって、アウトカムに関する短期～長期のロジック・モデルを描く場合は、純効果の測定可能性という点も考慮しなければならない。

アウトカムの測定手法は、こちらも社会科学の技法を応用して厳密に行われる必要があるが、その主たる原理は「比較」である。図2においても、プログラムの前後での比較、あるいは、プログラムを受けた状態とそうでない状態の比較という観点が入っている。もっとも正確にプログラムの純効果を把握するためには、プログラムを与える介入群とそうでない対照群をつくって、比較実験することである。これは無作為化フィールド実験法と呼ばれる。

一例を示すと²⁰、ニューヨーク市で1972年から1978年にかけて、薬物乱用者が小グループを組んで職業経験を積むという施策が施行され、その効果をはかるために実験が行われた。対象とされた薬物乱用者の条件は、18歳以上で過去3カ月以上の薬物乱用の治療経験があり、現在もそして過去24カ月の

うち 12 カ月以上は失業していて、公的な生活保護を受け、参加時に薬物中毒症状を起こしていない等である。自主的に実験に参加した者を、施策適用の 194 人と不適用の 207 人で、ほぼ同じ条件となる二つのグループに分けた。そのうちの一方のグループにある施策を施す。内容は、同じバックグラウンドをもつ者どうしで小規模チームを組み、市から支援サービスを受けながら共同で働く経験を積むということであった。割り当てられた典型的な仕事はオフィス事務、書類配達、ビルの清掃などで、指名されたチームリーダーを中心に易しい仕事から徐々に本格的な仕事に入り、金銭的報酬も次第に増やしていった。チームは厳しい指導を受けることもあるが、親身にあたってくれる支援者がいて、チームの中で日常的に話し合いももたれた。こうした共同労働の一定期間が終了すると、そこで身に付けた技術や経験を生かして、各自が自力で仕事を探した。もう一方のグループは、対照群として、特別なサービスを受けずにふつうに生活した後、仕事を探した。そして、両者の様子を 3 年間追跡して測定するわけである。

測定のための指標は、①3 年間で働いた週数、②収入額、③福祉依存率、④犯罪によって逮捕された率、⑤薬物・アルコール依存の比率、⑥結婚率と結婚生活の持続の程度などであった。3 年間の成果測定の結果、施策適用の有無による二つのグループで違いが大きかったのは、労働週数で、施策適用グループが 101 時間に対し、対照群が 46 時間と 2 倍以上の開きがあった。この他、年収で 2 倍以上、生活保護の受給率では約 8 分の 1、最初の年に逮捕された率で約 3 分の 2 など、施策適用のグループに生活改善の兆候がみられたのである。これによって、施策の効果がおおむね確認されたということで、この共同労働経験の施策は、他の地域にも適用されることとなった。

薬物依存者の更生プログラムを公民館でおこなっているところはほとんどないかもしれないが、あえて分かりやすい事例を示した。ここから、比較によってプログラムの純効果をはかるうえで、妥当な指標の設定であるとか、成果測定の厳密性、判断尺度の設定など、評価にあたってクリアしなければならない諸課題が見えてくるためである。これらの点について信頼できる手法を獲得

して、そこに、公民館ならではの機能や価値を反映させられれば応用範囲は広がる。公民館の活動においては、成果の測定において数値化されにくい部分や、自己評価か外部評価かということも含め、誰がどのような立ち位置で評価をするのかといった点も独自に考慮しなければならないだろう。

以上検討したアウトカム概念は、あくまで「プログラム評価論」の枠組みにおいて考えられたものであるから、最終的には実施したプログラムの価値をはかるために活用されなければならない。つまり、講座においてどれぐらいの知識を受講生に伝達できたかという点をアウトカムとして測定したい場合には、学習者が講座以外で学んだことについては、捨象しながら分析しなければならないということである。社会教育の活動においては、学習者の主体性が重視される。しかしながら、「学習者中心」という響きのいい価値観に浸りきって、本当は学習者が講座とは無関係に学んだ成果を講座の成果と混同し、ともすれば無意識のうちに学習活動を学習者へ丸投げする状態を引き起こしていることがないだろうか。この点は、公民館の立場からみるとよほど意識的に留意しておかなければプログラム評価論を活用する意義が無くなるようなポイントであるように思われる。

おわりに

アメリカで発展したプログラム評価論は、当該のプログラムがもつ効果を知りたいという動機から、データの収集や分析において社会科学の技法を応用しつつ、できるだけ外的要因を取り除いた「成果（アウトカム）」の形でプログラムの効果を抽出しようとした。投入された資源が成果として生成される道筋が見えなくなれば、プログラムの純効果を測ることもできなくなるから、実行されたプログラムの「プロセス」についても主要な評価対象となる。

評価の原理として注目されるのは、第一に、資源の投入から活動とその結果、そしてそれが成果につながる一連の過程を構造化していること、それに関連して第二に、具体的に評価したい事柄を評価クエスチョンとして設定し、評価計画に基づいた評価活動を行うということである。この二点はすぐにでも応用可

能な視点であると思われる。確かにこれまでのような事務事業評価シートの作成を中心に展開してきた公民館評価からみればなじみが薄いようにも感じるが、要は、評価の枠組みから関係者で決めていこうとするもので、ここにこそ、現状における公民館評価の弱点があると思われるのである。

プログラム評価について述べた諸論を読んでいると、たびたび「利害関係者」という文言に出くわす。アメリカではそれだけ利害関係者の意向は評価活動をあつらえる過程で無視できないものとされているのであるが、言い換えれば、関係者の多くが当該のプログラムによってもたらされる自分たちの利益に関心を持っていることのあらわれであろう。利害関係者のタイプによって設定する成果が異なるから、必要とされる評価の具体像も個別的で多様になるのが、プログラム評価論の特徴でもあった。

翻って、我が国の公民館をとりまく状況をみると、公民館に関わる利害関係者は、施設職員、有識者や街の名士による審議委員、さらに、利用者といっても住民の中で特定の層に偏っているなど、なんとも限定的で、ようやく学術的な検討に手をつけはじめた研究者を含めても、公民館にごく近い人たちだけで構成される傾向にある。利害関係者が固定化されれば、評価の多様性も望めない。そうであれば、新たな評価論を検討する内発的な動機はどこにあるのかということになる。

こうした場合には、公民館における活動を活性化させるためにも、まずは利害関係者として公民館に関心を寄せる層を開拓していくことが重要である。言い換えれば、これまで公民館に縁遠かった人や団体を利害関係者として巻き込んでいくということになるが、その際、プログラム評価論を呼び水とすることも可能ではないか。この理論の眼目はプログラムがもたらす成果の道筋を目に見えるように描くことにあるが、それによって、当該プログラムについて関係者間で議論されることが期待されている。多くの人たちで議論できるようになると、お決まりのように実施されてきた事業のあり方にも変化と躍動感が生まれるであろう。

最後に、本稿の問題意識に引きつけていけば、プログラム評価論は、公民館

が実施している活動の総体をひとつのコンテンツとしたときの価値について、プロセスからアウトカムまでを論理的に評価しようとするものと捉えるべきである。これは、学習者が何を獲得してどのように成長したのかという従来の学習論で中心的に論じられてきた側面とは焦点が異なるということに注意が必要である。誤解のないように申し添えれば、学習者中心のパースペクティブを否定するということではない。むしろ、自己主導的学習を展開する際、学習者にとっては眼前のコンテンツの様態がはっきりしなければ、それが自分にとってどんな価値をもつのか、はかりようがないということになる。よって、学習支援を充実させるためにも、プログラムの価値をきっちり評価しておくことが必要であるのだ。

その意味では、評価手法の体系化は、評価に対する信用性を高めるうえでとりわけ重要な課題である。定量的評価か定性的評価か、自己評価か第三者評価かといったことも、それだけを切り取って云々と論じてあまり意味をなさない。評価の「あつらえ方」によって、データ収集や判断基準の制定に適した手法があり、それをどこまでの条件で評価活動として実施できる資源をもつかといった検討を経て、はじめて、評価担当者や評価のタイミングの妥当性が決定できるのである。これらについては稿を改めて検討してみたい。

注

- 1 千葉県公民館研究委員会『さらに充実した公民館事業を目指すための自己点検評価に関する研究報告書』千葉県公民館連絡講義会、2007年。
福生市の取り組みについては、後掲3の伊東論文に詳しい。
- 2 『月刊公民館』（第一法規）
 - ・2004年9月号「特集 公民館の評価を考える」
 - ・2008年6月号「特集 評価で磨く！公民館」『月刊社会教育』（国土社）
 - ・2009年9月号「特集 私たちが求める公民館－評価づくり視点」
- 3 伊東静一「公民館事業評価に関する考察」『日本公民館学会年報』第6号、2009年、48-60頁。
茂木勇「公民館事業を伸ばす事業評価のあり方に関する研究」『日本地域政策研究』（8）、2010、309-316頁。

-
- 4 上野景三「公民館が果たしてきた役割と評価の視点」『月刊社会教育』国土社、2009年9月号、5-12頁。
- 5 Shadish, W. R., Luellen, J., "HISTORY OF EVALUATION", Mathison, S. (ed.), *Encyclopedia of Evaluation*, Sage Publications, 2005, pp.183-186.
- 6 小池源吾、佐々木保孝「公民館評価研究序説－「プログラム評価」の理論と手法－」広島大学大学院教育学研究科教育学教室『教育科学』第28号、2011年、25-41頁。
- 7 Knox, A.B., *Evaluation for Continuing Education, A Comprehensive Guide to Success*, Jossey-Bass, 2002.
- 8 龍慶昭、佐々木昭『政策評価 トレーニングブック』多賀出版、2003年、27-36頁。
- 9 Knowlton, L. W., Phillips, C. C., *The Logic Model Guidebook, Better Strategies for Great Results*, Sage Publications, 2009.
- 10 P.H.ロッシ、M.W.リップセイ、H.E.フリーマン著、大島巖、平岡公一、森俊夫、元永拓郎監訳『プログラム評価の理論と方法－システムティックな対人サービス・政策評価の実践ガイド』日本評論社、2005年、32頁。
- 11 同上書、31-63頁。
- 12 同上書、29頁。
- 13 同上書、48頁。
- 14 同上書、69頁。
- 15 Knowlton, L. W., Phillips, C. C., *op.cit.* p.36.
- 16 P.H.ロッシ他、前掲書、193頁。
- 17 同上
- 18 同上書、195頁。
- 19 「インパクト」のとらえ方について、ノールトンらが活動したケログ財団ではアウトプットとアウトカムの次の段階におこる長期的な効果とされているが、ロッシらは短期から長期にかけてのアウトカムの変化(すなわち、この部分のロジック・モデルそのもの)をインパクト・セオリーとして論を展開している。この点の説明については、安田節之、渡辺直登『プログラム評価研究の方法』新曜社、2008年、104頁を参照。
- 20 龍、佐々木、前掲書、16-17頁。
