



Title	評価と社会的インパクト：評価学の視点から
Author(s)	源, 由理子
Citation	第19回JINSHA 情報共有会：「社会的インパクト」を「評価学」の視点から考える, 11-38
Issue Date	2025-03-31
DOI	https://doi.org/10.14943/JINSHA-19.11
Doc URL	https://hdl.handle.net/2115/94735
Type	conference paper
File Information	011-038.pdf



評価と社会的インパクト ～評価学の視点から～

明治大学 公共政策大学院 ガバナンス研究科・教授 源 由理子

はじめに～評価学との出会い

ただいまご紹介いただきました明治大学公共政策大学院ガバナンス研究科の源です。本日はどうぞよろしくお願いいたします。

私は、評価学を研究しております。これまで「研究評価」については全く話したことはなく、研究評価の研究も行ったことがないので、私がお引き受けしていいのですかと申し上げたのですが、評価学の視点で最近いろいろ耳にする「社会的インパクト」はどのように捉えられるのかという話をしてほしいというご依頼でしたので、お受けすることにいたしました。この後、押海さんの方で研究評価についてのお話がありますので、そちらにつなげるような形で、評価学という視点に立ったときに、「社会的インパクト」をどのように考えることができるのかというお話をいたします。

今日はフロアの方に、評価学会から研究評価のご専門の方も来ているので、少しやりにくいところもあります。私自身は、評価の中でも、特に国際開発領域、社会福祉領域や文化芸術領域における評価の実践と研究を20年ぐらい行っています。

まず自己紹介をさせていただきます。私は大学を卒業してすぐJICAに就職し、国際協力の領域で仕事をしてまいりました。ご存じかもしれませんが、JICAや国際機関には、多額の評価予算がついています。おそらく今でも日本の中で評価のための予算が一番多く配分されていると思います。評価の予算といえば、コンサルタントへの委託料が主ですが、評価調査にかかる多額の予算がついていて、そういう意味で実務の中で評価に携わったのが、私の最初的评价学との出会いです。

評価調査に従事し始めた頃の私は、評価というのは測定できればいい、つまり、社会調査のノウハウがあればできるのではないかと、指標を測定し定量的に把握して統計分析するものではないかと、そういうイメージを持っていました。

ですが、ケニア・ナイロビのスラム地域での生活改善プロジェクト評価の仕事を担当した時—— JICA を辞めてからの仕事ですが——、資金提供者側が決めた指標だけで評価していいのだろうかという疑問が湧いてきました。貧困地域の生活が改善したという状態を評価するのに、資金提供者が決めた指標を測定するだけでいいのだろうかと思ったわけです。当事者である地域の人々、つまり調査対象となるコミュニティの人々がどのように考えているのかを考慮しなくていいのだろうか、ふと思いました。とても難しそうな評価だから、この仕事に従事することはやめようかと思ったくらいです。20年以上前のことですが、当時は日本には評価についての文献がほとんどなかったので、評価について調べてみようとしてアメリカの文献を読んだ時に、参加型評価というアプローチがあるのを見つけました。そのアプローチは、関係者つまり専門家も実施に関わった人たちも当事者も、みんなで何を価値あるものとして測るべきかを一緒に議論していく場を含んだ評価、対話を含んだ評価というアプローチでした。その時、これは、もしかすると面白いかもしれないと思いました。測定は評価調査の一部で、測定されたものをどう価値判断するのが評価だと定義したときに、その価値を決めるのは誰なのかという大きなテーマがあったからです。そんなところから評価にどんどのめり込んでいったという経緯になります。

最近取り組んでいる領域は、先ほど申し上げた社会福祉分野や芸術文化分野に加えて、政策評価や自治体・行政評価等になります。あと、社会的インパクト・マネジメントにも取り組んでおりますが、これが、今日お話しする社会的インパクトに関連する領域と重なると思います。

最初に申し上げておきますが、評価学において、社会的インパクトがこう捉えられるという共通の理解があるわけではありません。同じように混乱しています。ですから、今日私がお話しする内容は、私見も入っておりますが、評価学の基本に立ち戻ったときにどのようなところを考えていかなければならない

のか、もしかしたら新しい評価の領域として学会も含めて関係者と共に取り組む必要があるのではないかと感じております。そのあたりをお話したいと思います。

今日の構成です（スライド3ページ）。Part1「評価学における“評価”の考え方」で、主に「プログラム評価論」をご紹介します。後ほどご説明いたしますが、そもそも評価事業や、何らかの社会的介入の評価はどういう特徴があるのかが一点目となります。二点目は、評価学における「インパクト評価」という言葉と「社会的インパクト評価」の関係、評価学の視点で見たときに社会的インパクト評価はどのように捉えることができるのかということ、先ほど中野さんからインパクト投資の話がありましたが、そのあたりと合わせてお話できればと思います。最後に評価学の切り口から、研究評価について皆さんはどのように考えますかという問いかけで、まとめさせていただけたらと思います。

評価学における“評価”の考え方～プログラム評価の紹介

〈評価の定義と評価の目的〉

最初に評価学による評価の考え方です。実は、評価学という確立したものがあるかどうかは、心もとないところではあります。学会自体は、2000年に政策評価法が施行された時に設立された、比較的新しい500人くらいの小規模の学会で、その中にも様々な領域の方たちが入っておりますので、評価を横断的に捉えたとき、多様な議論ができる場ではないかと私は考えております。その中で、中心的方法論として一定の合意を得ている「プログラム評価論」、英語の Program Evaluation を訳していますが、こちらの方法を中心にお話しさせていただきます。評価といっても、人事とかプロダクトの評価とかいろいろありますが、このプログラム評価論は、特に社会的介入（social intervention）プログラムの評価で使われている方法論です。評価の方法論として、ある程度確立している領域ですし、今回のテーマとプログラム評価論の視点が良く合うのではないかと思います。「評価は理論と実践の融合である」と言われるように評価には実践が伴わないと全く意味がありません。そういう意味では、応用

社会科学であり、もっと言い換えると、評価のアプローチそのものが、社会の変化に応じて変化したり、新たなアプローチを提案したりすることになる等の特徴が強い領域だと考えております。

次に評価の原義についてです（スライド5ページ）。定義のようなものですので、ご存じかもしれません。とかく評価というと、私たちには成績表とかランキングとかそういったイメージがありますが、エバリュエイト（evaluate）はもともとラテン語で、eというのは引き出すという意味、バリュー（value）は価値だと考えると、評価とは評価対象の価値を引き出す行為であるとよく言われます。つまり、評価という道具を使って、取り組んでいることの価値を高めていくことが評価そもそもの目的ということになります。先ほど冒頭で申し上げましたように、私はケニアのプロジェクト評価に関わっていましたが、プロジェクト評価に関わる前は、指標を測定する、つまりデータを収集していくような社会調査ができればいいのだと思っていた節がありました。もちろん、指標の測定も非常に重要なのですが、それはイコール評価ではない。評価とは、事実の特定にその価値判断が加わるということだと言えます。

例えば、何らかの状況が50%改善したという事実を、50%も改善したと言ったり、50%しか改善しなかったと言ったりすることがあります。これは誰の判断でしょうか。そういう意味では、評価行為は客観的ではないかもしれません。データはできるだけ客観的に集めるべきだと思いますが、それをどう価値判断するのか、どういう合意形成のもと判断していくのかが、非常に重要になってきます。

「評価は物事のメリット、値打ち、意義を体系的に明らかにする」と定義されます。「メリット」というのは、そもそも実施していることが内在的に持っている本質的な価値と捉えることができます。「値打ち」は、貨幣のような共通のものさしを使って、共有の価値として持っているものです。貨幣は価値基準の一つで、他にも基準になるものはあると思います。「意義」は、誰にとって重要なのか、実施組織のミッションにとってどういう意味を持つのかになります。価値といっても、メリット、値打ち、意義といった複数の視点から捉えていく必要があります。

プログラム評価論というプログラムは、社会的介入、つまり、何らかの社会的目的を持つ事業やプロジェクト、行政であれば政策や施策、そういったものを指します。また何らかの社会的課題解決や社会的価値創造を目的とした事業の評価も、プログラム評価で体系的に評価することができます。評価の方法は、社会調査手法等を使ってプログラムの有効性を体系的に評価するのですが、各プログラムを取り巻く政治的な要因や社会的要因は多様ですので、一つの方法が全てに適用できる訳ではありません。その意味で、評価のアプローチは、評価の目的や評価結果の利用者に応じてカスタマイズされるべきだとされています。テラーメイドという言葉を使っていますが、評価対象となるプログラムの文脈に合わせて評価のアプローチを考えないと評価結果は使われないものになってしまいます。評価結果の活用とは、社会状況を改善していくための社会的活動の情報源になるということであり、それがプログラム評価の目的であります。そうであれば、使われないものを発信しても意味がなく、また評価に使ったお金が無駄になります。こういった点が、プログラム評価の大きな特徴となります。

プログラム評価論というところの評価の特性について、もう少しお話しさせていただきます。評価の活用目的には大きく分けて三つあると言われています(スライド7ページ)。

まず一つ目は、事業の評価結果を事業の改善に使っていく、あるいは何らかの変革に使っていくというものです。改善は今あるものを良くしていくというイメージですが、変革とはトランスフォーメーションですから、全く違う形で新しく目指すものというイメージです。二つ目がアカウンタビリティ、つまり説明責任の確保、三つ目は新たな知識の創造です。

一点目の改革・改善という意味で重要になるのが、インパクトや成果よりも、どちらかという実施のプロセスだと思っています。活動を実施しているプロセスそのものは妥当なのか、本当にこの方法でいいのかを評価し検討します。

二点目のアカウンタビリティで、理論的に言われるのは、資金提供者に対する説明責任です。もらったお金を使ってこれだけの成果がきちんと出せたこと

を説明する責任となります。ですが、目標が達成されたと言ったとき、そもその目標設定が誤りだったらどうなるでしょう。そういう意味では、私はアカウンタビリティには、リスポンシビリティ（responsibility）という概念も入ってくるのではないかと考えています。これは、目指すことが妥当なのか正しいのかという実施内容に対する責任を意味しています。またアカウンタビリティなので、誰に対しての説明責任なのかの検討も必要です。資金提供者なのか、あるいは地域や社会の人々なのか、また何らかの課題解決という意味では、その課題を抱えている当事者に対する説明責任なのか、社会全体に対する説明責任なのか、それによって評価でどういうところを重点的に見ていくのかも違います。発信の仕方も、誰に向けて説明責任を果たしていくのかによって異なると思います。

それから三点目として新たな知識創造です。これは、評価研究という分野になると思います。評価をすることによって新たな知見を発見するという実証研究を踏まえた取り組みというのは、このジャンルに入るのではないかと思います。

この三つが独立して存在するというよりも、互いに影響しあっているため、境界線を引くのは結構難しいと思います。この評価がどのように、誰に活用されることを目的に行われているのかということは、評価結果を社会状況の改善に使うという意味ではとても重要なポイントになってくると思います。

さらにその目的によって評価主体が分かれてきます。内部評価や自己評価もあれば、外部評価、第三者評価もあり、関係者とともに評価を行うこともあります。冒頭で申し上げたケニアの事例のように、多様な関係者が評価を行う過程で何らかの価値を相互の協働の中で生み出していく、そういうプロセスを含む評価のアプローチもあります。

このようなプログラム評価の三つの特性から見る「評価」の要諦・特徴が三点あります（スライド9ページ）。三点目は今申し上げた、評価の目的によってアプローチが異なるということですが、一点目の「結果や成果を見ることだけが評価ではない」とは、実施プロセスも重要になってくるといことです。二点目が「評価対象であるプログラムの構造を明らかにする」ですが、自分た

ちが何を評価対象にしているかがわからないと、どういう評価が必要なのかもわからないという意味になります。

これらの特徴に関連して、プログラム評価の中心的な考え方が二つあります。一つ目は、スライド9ページの一点目に挙げている「評価5階層」という考え方で、決してインパクトや結果・成果だけではなく、プロセスも重要であるという点です。この後、ご説明します。二つ目に、評価を行う前に評価対象が評価可能なかどうかを考えなくてはならないということ、つまり Evaluability Assessment だと思っています。そして、どのような構造で何を狙いとして、どのような目的でどんなことを実施しているのかを検討するために、ロジックモデルや変化の理論を道具として使用するということになります。これらについて触れたいと思います。

＜“プログラム評価”の複合的視点：評価は結果や成果をみることだけではない＞

まず、複合的な評価の視点です（スライド10ページ）。評価は結果だけを見るものではないという事例の一つとして、あるNPOの生活困窮世帯の子どもへの学習支援があった場合に、学習支援を行うと子どもの学校での成績が上がるという仮説を立てて活動しているとします。皆さん、これをどう思われますか。本当に目的はこれでいいのか。学校の成績が上がることを目的にしているのか。仮に、子供の学校での成績が上がることを目的に置いたとして、学習支援はこのやり方だけで十分なのか。これらは、戦略に関わる問いです。この設計でいいんだろうかという視点です。あるいは実施プロセスでは本当にきちんと支援が行われているのだろうか。これはプロセスの妥当性を問う視点です。このように、事業の評価を求められたときには、複数の視点のクエスチョンが出てくる可能性があります。

プログラム評価では、五つの評価の視点があると言われています。スライド11ページの一番下からです。まず、ニーズ評価です。そもそも事業が目指そうとしていることはこれで妥当なのかを評価します。非常に重要なポイントですが、実際の事業では十分に把握されていないことが多いところです。またこれは事業を実施する前の評価です。次にセオリー評価では、このやり方で成果

を上げることができるだろうか、本当に事業設計は妥当なのだろうか、などを評価します。妥当性の中には、倫理的な視点も含めて、持続可能性など様々な視点が入ってくると思います。三点目が実施プロセスの評価です。そして四点目にプログラムのアウトカム、つまりインパクト評価となります。インパクトやアウトカムを社会課題解決型で捉えたときには、当然ニーズ評価と呼応します。これが問題だから、こういったプログラムを実施するのですが、インパクト評価では当初のニーズは満たされているかが評価される。最後にもちろんお金、効率性です。投入した資源に見合っってこのような結果が出ているのかどうかを評価します。このように複数の視点で見ることになります。

このような「評価の複合的視点」が意味するところを、私は次のように考えています。どのような指標で成果を見るのかの前に、そもそもどのような社会課題を解決しようとしているのか、あるいはどのような社会的価値を創出しようとしているかの問いかけがまず必要だということです。また、インパクトやアウトカムの達成は、一つの介入で可能なのかという問題もあります。ここはおそらく社会的インパクトと関係すると思いますが、複数のステークホルダーが関係し合うことで生み出されるものが多いと思います。それは社会が多様なステークホルダーで動いているからですが、それを一点の介入だけで達成するのは難しいと考えます。そういった意味で、関係者間の対話や討議・熟議を通して、何をもってよしとするのかという議論が必要だと考えています。一回の「評価調査」ではなく、継続的に評価の視点を持つということが、5階層という複数の視点から評価をする意義になります。

〈評価ツールとしてのロジックモデルの特徴：アウトカム志向とプロセス重視〉

二点目がロジックモデルになります。これは、皆さんご存じのことだと思います。インプット、活動、アウトプット、アウトカムという構成要素で捉えています。ここで重要なのは、インプットからアウトプットまでがプロセス理論と呼ばれていて、マネジメント側がコントロールできるとされている点です。もう一つ、インパクト理論がありますが、ここではこういう活動をしたらインパクトが生まれるだろうという仮説を立てます。このインパクトは、

インパクト理論の言うところのインパクトですが、これはロジックモデルそのものの言葉ではアウトカムとなります。ですので、アウトカムを評価するといった場合は、アウトカムの測定ではなく、評価をするという意味で、インパクト評価と私はイコールだと思っています。

そこで重要となるのが、帰属性です。本当にこの介入のおかげでインパクトが生れたのか、他の要因でないのかを分析します。当然、インパクトは介入以外の要素がとて多いわけです。これはどんな小さな介入でもおそらくそうです。ですので、この点がこの後、社会的インパクトを見るときにキーになる部分だと思います。つまり、インパクトは外部要因の影響が大きいということになります。スライド 15 ページをご覧ください。実は先ほど事例としてお示した、学校の成績が上がることを目的にするという NPO は、実際には、子供が勉強を楽しみ、希望を持って自分の進路に歩み出せるようになるという目的で取り組んでおり、これが実施組織である NPO のミッションとなっています。つまりこの NPO は、子どもたちの自己肯定感が低いことが問題だから、子ども支援に取り組んでいる。そうであれば、プロセス理論に示された介入だけでは不十分で、より帰属性を高めるために、スライド 16 ページの縦のロジックモデルを見ていただくとわかるかと思いますが、保護者への働きかけや、地域社会、様々なステークホルダーへの働きかけがなければ、子どもたちは地域の資源とつながりながら自分たちで立ち向かい、希望の進路に歩み出せるようにはならないだろうということです。ロジックモデルというのは、プロセスを見直していくための共通の言語のような働きがあります。また、その帰属性の部分が評価の際に非常に重要になってくるということでもあります。

ただし、こういった「プログラムの構造化」を考えたときに「手段—目的」の構造をあらかじめ推論することが全てのプログラムで可能かどうかという課題に直面します。これは現在、芸術文化領域の評価を行っていて、本当に実感しています。ロジックモデルをきれいに最初から書くエビデンスがあれば別ですが、そもそも取り組んでいることの特性を捉えたときに、あらかじめ目的を設定すること自体がその介入の本質的価値を削ぐかもしれないという、そういう問題が出てくるわけです。これはおそらく社会的インパクトを考えるときも

関係するのではないかと思います。

二点目に書きましたように、介入の「本質的価値」を内在的価値としたとします。それから、「手段的価値」、これは外在的価値です。例えば、地域のアートプロジェクトでは、多様なアートの試みを介して、そこにいる人々や一緒にクリエーションプロセスに参加した人たちが、自分たちの多様な価値との関わりや、遭遇から生まれる新たな価値を生み出していきます。これはアートだから生まれる本質的な価値だと捉えることができると思います。目的として、例えば、アートを通して多様なルーツを持つ人たちとの共生文化が育まれるという社会的な目的、公共的な価値のようなものがあつたとします。こういう目的を前提に助成金が付くとします。そのようなプログラムを評価する際に非常に難しいのが、この帰属性です。アートに接する、アートに参加する、アートであるからこそ生まれる価値という本質的なものと、それが何らかの社会課題解決に結びつく。ですが、この間にある結びつきについて、いわゆるエビデンスを示して介入を実証的に見ていくのは非常に難しい。なぜなら、そこには様々な外在的要因の影響があるからです。ではどうしたらいいのか、という点は今模索をしているところでもあります。

「インパクト評価」と「社会的インパクト評価」～新たな評価の可能性?～

〈評価学でいう「インパクト評価」：「帰属性」との関係〉

「インパクト評価」と「社会的インパクト評価」ということで、今申し上げたことを踏まえて私自身が考えているのは、新たな評価の可能性を評価学の中で考えていくことができるのではないかということです。

評価学におけるインパクト評価は、いわゆる RCT というランダム化比較試験で検証されたりします。RCT では、プログラムに参加した人と参加していない人を比べて、純粋な効果をインパクトと捉えるのがいわゆるインパクト評価のゴールデングルールだと言われています。統制群を立てて行うもので、最初からランダムにプログラムに参加する人と参加していない人を作って、その効果を比較するという方法です。介入したグループと介入のないグループとの差が本当のインパクトだという捉え方です。統制群がないと、観測される効果が

もっと大きくなるわけです。ですが、そのプロジェクトに参加していない人も何らかの影響を受けていますので、参加した人と参加しない人の変化の差が「インパクト」になります（スライド 20 ページ）。

こういう考え方がゴールデンルールとしてある中で、この後押海さんのご発表でも出ると思いますが、エビデンスの序列というものがあります（スライド 21 ページ）。一番上が体系的レビュー、メタ分析、つまりランダム化比較試験したものを複数メタ分析するやり方です。これこそが、科学的なエビデンスだとして、序列の 1 位に位置しており、帰属性を証明する最も科学的な方法だという考え方とされています。そして、この下の順位にいろいろな方法論があります。下の方にポリシーブリーフや専門家の意見といったものや、事例研究などの極めて質的なデータによる評価があります。

私がエビデンスの序列のタイトルにクエスチョンマークをつけたのは、本当にこの序列という捉え方でいいのだろうかという一つの大きな疑問があるからです。それは先ほど申し上げたように、介入によっては帰属性を量的データだけで捉えることが難しいものがあります。あらかじめ量的データで指標を設定し目標値を設定した場合に、本来の介入の本質が歪められるものがあるため、質的なデータも非常に重要になってくると思っています。このエビデンスの序列の出自が医療領域であることを踏まえると、多様なステークホルダーの影響が考えられる社会的介入の領域にそのまま適用していいのだろうかとも考えます。

こういった疑問を持ちながら、一方で出てきたのが「社会的インパクト」という言葉になります。これは趣旨説明で中野さんの方からご紹介ありましたように、民間セクターの動きとして、インパクト投資や ESG 投資の中から出てきています。もともとは環境問題などのマイナスの影響が出た場合に投資を引き上げるというダイベストメントという動きで、1920 年代からすでに現れています。1980 年代からは、国際 NGO 等による環境活動でのネガティブスクリーニング、例えば武器輸出をしている、採掘による影響があるなど、環境に悪い影響を与える企業に対しては投資をしないという動きが出てきております。この背景には国連などが中心になって、様々な動きがあったということも

あります。それを受けて、社会的責任投資という考え方が出てきます。世界的に一番大きな資産運用会社が、社会にネガティブな影響を与えるものに対しては投資をしないという決定をしたということも投資行動に影響を与えたと思います。ESG投資や社会的インパクト投資では、経済的リターンと社会的リターン、この両方を見なければならいとされています。それぞれの割合の程度は違いますが、基本的にその両方を、しかも民間セクターがCSR（企業の社会的責任）として実施するというよりも、本業として社会的課題解決に参入していくという動きが出てきました。その中で生まれた言葉が実は「インパクト」ということになります。ここで言うインパクトは、評価学のインパクト評価とは出自が全く違い、資本市場での活動の成否を判断するものさしであるということになります。こういった正負の社会的、環境的価値、つまり、これが社会的インパクトですが、これを見える化する必要性が出てきたということになります。先ほどご紹介がありました内閣府や、他にもいろいろな文献で定義がなされています。

スライド24ページにあります、Epstein & Yuthasによる「活動や投資によって生み出される社会的・環境的变化」という定義は、投資を背景にして出てきていると思います。また、インパクト測定マネジメント（IMM）については、GIINという世界的な投資ネットワークが、効果的なインパクト投資を行うために不可欠であり、マイナスの影響はできるだけ軽減してプラス面を増やすための方法が含まれる本質的に反復的なものだ、と言っています。これについて、私は「社会的インパクト評価」というのは、どちらかという「社会的インパクト・マネジメント」だと捉えております。投資家と企業は、できるだけネガティブな面を出さないようにPDCAを回し、社会的リターンを増やしていく。かつ、もちろん経済的リターンも十分出していくということを、投資家が投資する前から企業と一緒にマネジメントしていく。そういう動きになります。

あと、投資家の視点として、社会的インパクトを生み出すためには、複数の企業の動きを見る必要があります。一つの企業の働きだけで社会的インパクトを生み出すのは難しい。先ほどもロジックモデルのところで申し上げました

が、社会的影響を生み出していくためには、おそらく一つの介入だけでは難しく、多様なステークホルダーが関わる、あるいは、ここで言えば多様な企業に関わるポートフォリオを前提としてインパクトを見ていくという動きになります。

そういった意味で、評価学の視点から「社会的インパクト評価」を見たときに、私は三つぐらいの特徴があると思っています。

まず一つ目に、「社会的インパクト評価」というのは、どちらかという社会、あるいは外部からの要請であるだろうということです。そういった意味では、評価の目的としては、アカウンタビリティの側面が強いのではないかと考えています。先ほど社会的責任投資という言葉がありましたが、投資家と事業実施者との関係性を築き上げなければ持続可能な社会とならない、という中で論じられてきたのが社会的責任投資です。どちらかという投資家、いわゆる機関投資家の視点がかなり強いです。機関投資家が投資してくれなければ企業は成長できません。両者は非常に重要なパートナーでありつつ、その中で社会的責任を果たすことになります。この点が、今回、人社系 URA ネットワークの皆さんから教えてもらった「責任ある研究・イノベーション (RRI)」という考え方と似ていると思っています。文献を読んで、似ているところがあるなと思いました。つまり、アカウンタビリティというのは、先ほど申し上げたように、資金提供者に対する説明責任だけではないので、誰に対するアカウンタビリティなのかを考える必要があるのではないかと考えています。

二点目が、「社会的インパクト評価」とは、評価学でいうところの「インパクト評価」とは必ずしも同じではないということです。私も内閣府の社会的インパクト評価検討ワーキング・グループに入っていましたが、当時、Impact Measurement and Management を「社会的インパクト評価」と訳した時、評価学会から反対がありました。これはインパクト評価ではない、エビデンスをきちんと出していくインパクト評価とは違うので、違う評価だ、という意見が出ました。確かにそうです。先ほどの定義を使うと違う。どちらかというマネジメントで、できるだけ負のインパクトをもたらしさないようなマネジメントをしているか、ミッションを踏まえた妥当な戦略形成がされているかどうかを

評価する。これは、プログラム評価の5階層でいうとセオリー評価ですが、そういった視点が強いかもしれません。

三点目が、参加型とか協働型という評価アプローチとの類似性です。つまり多様なステークホルダーとの協働プロセスから生み出される「価値」を創出し、可視化していく——そういう動きは、先ほど少し紹介したステークホルダーとともに評価をしていくというアプローチに似ているところがあるかもしれないという点です。そして、関係者間の合意形成をとおした推論です。これは仮にエビデンスを量的に見るのが実証主義に基づく研究だとすれば、私は構成主義的なアプローチだと思います。つまり現状どのような価値が生まれていて、それをどう可視化してどう積み上げて、見せていくか——質的で構成主義的なアプローチが「社会的インパクト評価」にあることを検討する余地があるかもしれないと考えています。

最後に、研究の「評価」への問いかけをして、押海さんにおつなぎいたします。評価学の視点から改めて問い直したときに、研究評価は、誰が評価主体で、何を評価対象とするのが問題になると思います。大学評価なのか研究評価なのか、大学のミッションにおける位置づけはどうか、あるいは活用目的は何で、誰がその評価結果の潜在的な利用者になるのか、です。評価学では、この、誰が潜在的な利用者になるのかを、評価の最初に考えます。そうでなければ評価しても無駄かもしれないからです。誰が使うのかは、どのように評価するのかの検討につながります。最後に、“測りすぎ”の弊害とスライドに書きましたが、研究評価だけではなく、私たちが評価をするときにいつも頭に入れておかなければならないものだと思っています。測りすぎて指標が一人歩きし、測定されることだけに注目が集まり、本質的な価値を阻害してしまうという部分を、本当に戒めなければならないと考えています。

以上が私からです。ご清聴ありがとうございました。

JINSHA情報共有会@北海道大学

評価と社会的インパクト ～評価学の視点から～

源 由理子

明治大学公共政策大学院ガバナンス研究科

2024年10月9日

・ 専門：評価論、社会開発論

- ・ わたしが「評価」に興味をもったきっかけ
 - ケニア国ナイロビ市スラム地域の生活改善プロジェクト
 - 既に決められている「資金提供者側」の指標で測っていいのだろうか、という疑問
 - 「参加型評価」への興味

・ 評価はおもしろい・・・かも

(最近取り組んでいる領域)

- ・ 社会福祉分野
- ・ 芸術文化分野
- ・ 政策評価、自治体行政評価
- ・ NPO事業評価、企業CSR評価
- ・ 社会的インパクト・マネジメント、等

自己紹介

本日の構成：

Part I: 評価学における“評価”の考え方：「プログラム評価」の紹介

- 評価の定義と評価の目的
- “プログラム評価”の複合的視点：評価は結果や成果を見ることだけではない
- 評価ツールとしてのロジックモデルの特徴：アウトカム志向とプロセス重視

Part II: 「インパクト評価」と「社会的インパクト評価」

- 評価学でいう「インパクト評価」：「帰属性」との関係
- 「社会的インパクト評価」の出自：IMMから考える

最後に：「研究評価」への問いかけ

©2024 Yuriko Minamoto. All rights reserved.

3



Part I: 評価学における“評価”の考え方：プログラム評価の紹介

“評価は理論と実践の融合である”

©2024 Yuriko Minamoto. All rights reserved.

4

評価の原義

e + value + ate

“評価は評価対象の価値を引き出すものである”

評価 = 事実特定 + 価値判断
指標の測定 ⇔ 評価

製品、人事、組織、事業などいろいろ

- ・「評価は物事のメリット、値打ち、意義を体系的に明らかにすることである」

(Scriven 1991)

“Evaluation is the systematic determination of merit, worth and significance of things”

©2024 Yuriko Minamoto. All rights reserved.

5

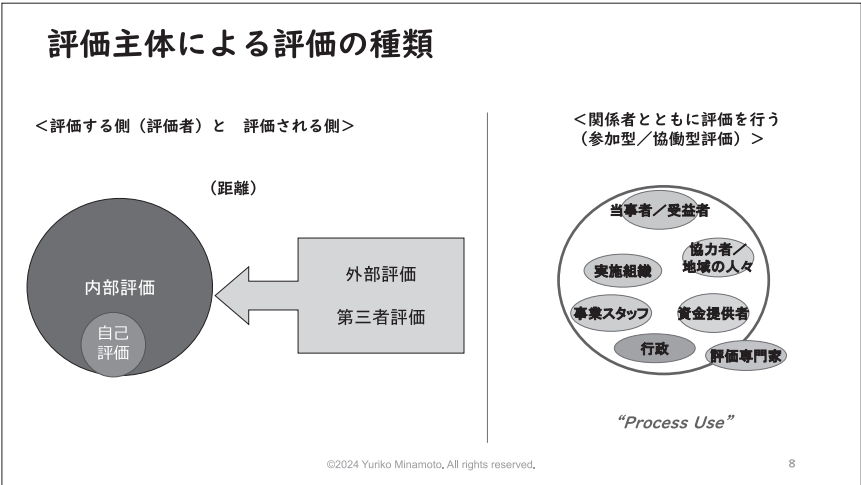
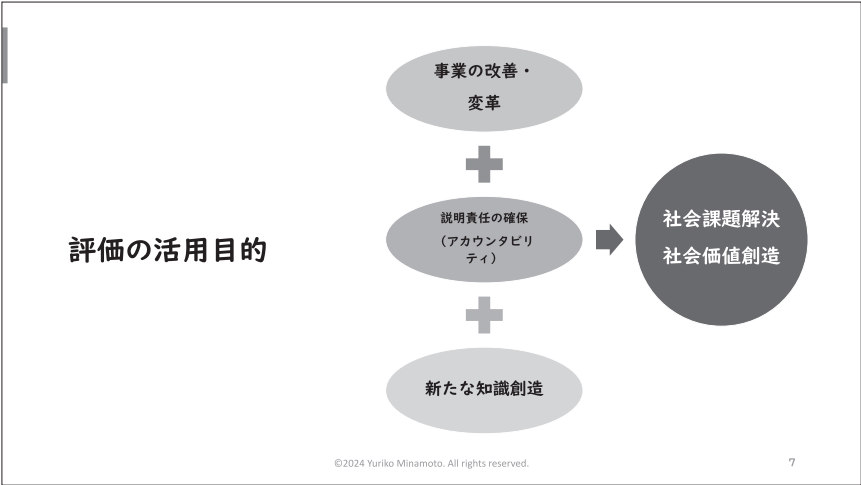
「プログラム評価 (Program Evaluation)」の方法論について

[出典：Rossi, Lipsey & Freeman 2004:29]

- ・プログラム：
 - ある社会的目的をもつ介入（政策、施策、事業、プロジェクト、活動群等）
 - 何らかの社会的課題解決、社会的価値創造を目的とした事業
- ・評価の方法：
 - 社会調査手法を使い、プログラムの有効性を体系的に評価
 - プログラムを取り巻く文脈を考慮して行うもの（評価のカスタマイズ）
- ・評価結果の活用：
 - 社会状況を改善していくための社会的活動の情報源となるものを提供

©2024 Yuriko Minamoto. All rights reserved.

6



プログラム評価の三つの特徴から見る「評価」の要諦

1. 結果や成果（インパクト）を見ることだけが評価ではない

- 評価の複合的視点：「評価5階層」
- インパクト志向とプロセス重視

2. 評価対象であるプログラムの構造を明らかにする

- 評価可能性調査（Evaluability Assessment）
- インパクト理論とプロセス理論（プログラムセオリー）の確認
- ロジックモデル、変化の理論（TOC: Theory of Change）など

3. 評価の目的によってアプローチが異なる

- 新たな知の発見か、外部統制への対応か（説明責任）、自己マネジメントのためか（改善・変革）
- 誰が評価するのか（評価主体）

© 2024 Yuriko Minamoto. All rights reserved.

9

1. 複合的な評価の視点

（例）あるNPOの生活困窮世帯の子どもへの学習支援

社会的なインパクト

活動の実施中に直面した困難はないの？どう対応したの？

そもそもめざすことはこれでもいいの？何を「良いコト」としてとらえているのか？

生活困窮世帯の子どもへ学習支援を行う

（目的・目指す成果：アウトカム）

子どもの学校での成績が上がる

この活動だけで目的は達成できるの？

© 2024 Yuriko Minamoto. All rights reserved.

10

プログラム評価の5階層 (Rossi, Lipsey and Freeman, 2004: 80)

〔評価は結果や成果をみることだけではない〕

プログラムの効率性評価:

投入した資源に見合っているだろうか?

プログラムのアウトカム/インパクト評価:

成果はあがったのだろうか/ニーズは満たされただろうか?

プログラムのプロセスと実施の評価 (プロセス評価):

実施過程に問題はないか?

プログラムのデザインと理論の評価 (セオリー評価):

このデザインで成果をあげることができるのだろうか? 事業の設計は妥当か?

プログラムのニーズ評価:

事業がめざそうとしていること (社会的ニーズ、社会的に生み出したい価値、ターゲットグループなど) は妥当か?



©2024 Yuriko Minamoto. All rights reserved.

11

“評価の複合的視点”が意味するところ



- どのような指標で成果を見るのかの前に、そもそも「どんな社会課題を解決しようとしているのか」、「どんな社会的価値を創出しようとしているのか」の問いかけがまず必要
- 社会的介入によるインパクト (アウトカムの達成) は、複数のステークホルダーが関係し合うことで生み出されるもので、関係者間の対話・熟議をとおして、何をもち「良しとするのか (評価基準)」の議論がまず必要

©2024 Yuriko Minamoto. All rights reserved.

12

(参考)

“評価的思考”で次のアクションを起こす

評価的思考 (evaluative thinking) とは、

- ・ 物事が起こる道筋が見える化し、
- ・ 意味のある問いかけを投げかけ、
- ・ 振り返り (省察) や、異なる視点に立って、物事を深く理解しようとし、
- ・ 情報に基づいた決断を下し、行動を起こす、

行動・認知プロセスである。

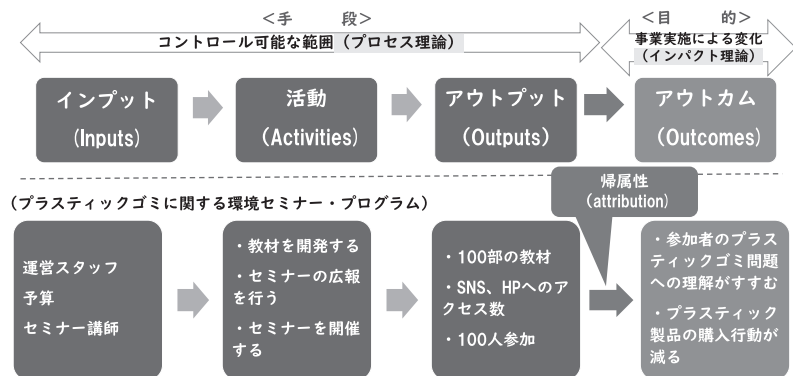
(出典：Archibald & Buckley, 2015をもとに源訳)

©2024 Yuriko Minamoto. All rights reserved.

13

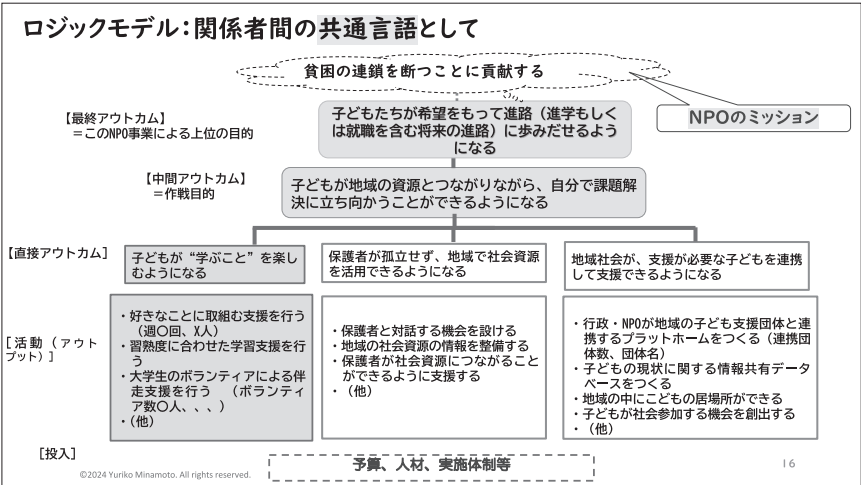
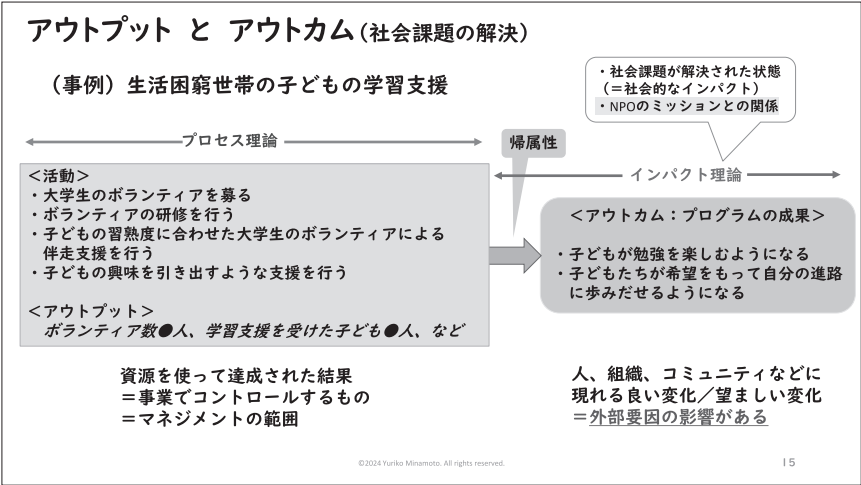
2. ロジックモデルの構成要素

出所：源（2020：44）図2-6に一部加筆



©2024 Yuriko Minamoto. All rights reserved.

14



“プログラムの構造化”を考えるとときに直面する課題

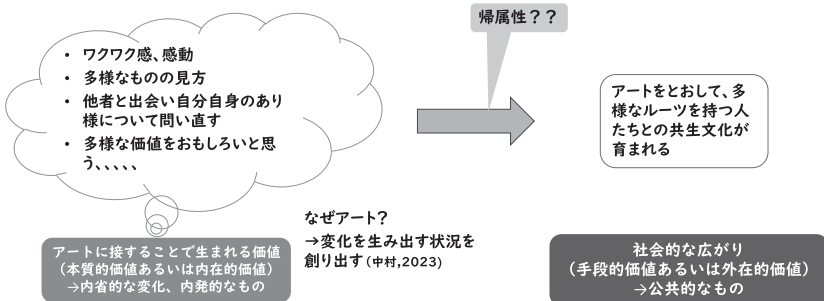


- 「手段-目的」の構造をあらかじめ推論することが、すべてのプログラムで可能か？
- 「介入の本質的価値（あるいは内在的価値）」と「手段的価値（あるいは外在的価値）」という考え方
（例：芸術文化領域）

©2024 Yuriko Minamoto. All rights reserved.

17

芸術文化領域における議論



©2024 Yuriko Minamoto. All rights reserved.

18



Part II: 「インパクト評価」と「社会的インパクト評価」 ～新たな評価の可能性?～

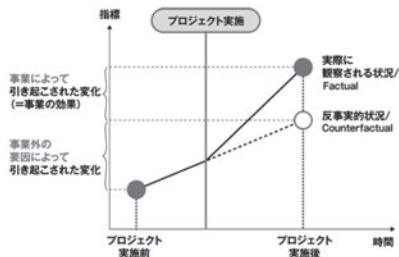
©2024 Yurika Minamoto. All rights reserved.

19

評価学における「インパクト」と「インパクト評価」について

インパクト評価概念図：

実際に観察される状況と反事実の状況との比較



インパクト：

介入（事業）による変化から、介入がなかった場合の変化（反事実）を差し引いたもの／純効果

RCT（ランダム化比較試験）等で因果関係の確からしさを検証する評価方法が最も科学的とされる

Methodology driven

出典：JICAインパクト評価サイト <https://www.jica.go.jp/activities/evaluation/>

©2024 Yurika Minamoto. All rights reserved.

20

エビデンスの序列???

エビデンス＝ある介入行為に対する因果的な検証結果＝インパクト評価の結果

順位	方法論	
1	体系的レビュー、メタ分析	新薬の有効性を示すエビデンス／福祉領域 の対人サービス分野のエビデンス等 量的データ／ 統計分析
2	ランダム化比較試験（RCT）	
3	コホート研究	
4	時系列研究	
5	統制群のある事例比較	質的データ／ 質的研究
6	横断研究	
7	事例研究、質的評価	
8	経済評価	
9	語りのレビュー	
10	ポリシーブリーフ、専門家の意見	

©2024 Yuriko Minamoto. All rights reserved.

出所：Crossick & Kaszynska（2022:266） 21

“科学のエビデンス”がすべてなのか？

～評価研究から政策評価への流れ～

実用重視型評価の登場（Theory Driven）1970's～

- 活用される評価への関心
- 実施プロセスを見ることも必要
- ロジックモデルによりプロセス理論とインパクト理論を明らかにしていく

文脈による違い

- 一般化が必ずしも目的ではない
- 実証研究から、質的研究の要素を入れた評価へのシフト

参加型、ピアレビュー、専門家による討議・熟議

- 何をもって「価値あること」と判断するのか

©2024 Yuriko Minamoto. All rights reserved.

22

社会的インパクト：民間セクターの動き

*出所：今田（2024）を参照し作成

企業活動から派生する負の影響（環境問題等）を低減するためのダイベストメント（1920's～）

国際NGO等によるネガティブスクリーニング（1980's～）



社会的責任投資（SRI）、ESG投資、インパクト投資

「経済的リターン」 vs. 「社会的リターン」



民間セクターが本業として社会課題解決に参入

“インパクト”：資本市場において活動の成否を判断するものさし



正負の「社会的・環境的価値」（＝社会的インパクト）を見える化する必要性

©2024 Yuriko Minamoto. All rights reserved.

23

「社会的インパクト評価」という言葉の登場：投資判断としての要請

・社会的インパクトの定義

- － 「活動や投資によって生み出される社会的・環境的变化」（Epstein & Yuthas, 2014）
- － 「短期・長期の変化を含め、当該事業や活動の結果として生じた社会的・環境的なアウトカム」（内閣府、2016）

・IMM（Impact Measurement and Management:インパクト測定・マネジメント）

- － GIIN（グローバル・インパクト投資ネットワーク）
- － 「効果的なインパクト投資を行うために不可欠で、自らの投資アプローチが人々と地球に与えプラスとマイナスの影響を特定し、特定の目標にそってマイナス面を軽減し、プラス面を最大化するための方法が含まれる」「本質的に反復的（iterative by nature）」である。



➤ 二つの動き

- ① 投資家と企業によりPDCAサイクルを回す
- ② 投資家の視点：社会的インパクトを生み出すためには複数の企業を動きを俯瞰したポートフォリオレベルのインパクトを見る*

©2024 Yuriko Minamoto. All rights reserved.

*出所：林（2024）

24

評価学と「社会的インパクト評価」

外部からの要請 => アカウンタビリティ	「インパクト評価」とは 必ずしも同じではない	「参加型／協働型評価」 のアプローチとの類似性
- 投資家と事業実施者との関係性を築きあげなければ、持続可能な社会とならない（社会的責任投資：SRI） - 責任ある研究・イノベーション（RRI）との類似性*	- マネジメントの関与：できるだけ負のインパクトをもたらしさないようなマネジメントに重きをおいている - ミッションを踏まえた妥当な「戦略形成」（セオリー評価）の視点	- 複数のステークホルダーとの協働プロセスから生み出される“価値”の創出と可視化 - 関係者間の合意形成（納得性）をとおした推論

*参考：吉澤（2013）

©2024 Yuriko Minamoto. All rights reserved.

25

最後に：研究の「評価」への問いかけです

- 誰が評価主体で、何を評価対象（評価対象の構造）とするのか
 - 大学のミッションにおける位置づけ？
 - なぜ「研究評価」なのか
- 活用目的は何で、誰がその評価結果の潜在的利用者（potential user）なのか
 - 誰が評価結果を使い、何に活かすのか
- どのように評価するのか
 - 研究が生み出す社会的・環境的な「価値」をどのように言語化・可視化するのか
 - “測りすぎ” * の弊害：本質的な価値との齟齬を生み出す危険性

* Muller（2018）

©2024 Yuriko Minamoto. All rights reserved.

26

(参考) 社会的介入に関する評価論の系譜

出所: 源 (2016) を基に加筆・作成

評価目的・文脈ごとに方法論を選択する

I. 1960's- : 科学的評価の主流化

- ‘科学的に正しい’ 根拠 (エビデンス) を提供することが評価の役割 → EBM (Evidence Based Medicine)
- 実験デザイン手法等による「介入」と「成果」の帰属性を問う

- Truth ‘科学的に実証された真実’

- Methodology driven
- 定量評価/インパクト評価/外部評価

II. 1970'- : 実用重視の評価

- ‘役に立つ’ 知識を提供、あるいはその活用を促すことが評価の役割 (プログラム評価、参加型評価という言葉の登場)
- 政策評価の領域からの問題提起

- Use 実用、活用

- Theory driven
- 定性評価の導入/プロセス評価
- 内部評価、参加型評価

III. 1980's- : 不確実性に対応した評価実践の提案

- 評価の目的やおかれた状況によって評価の手法は異なる
- 評価方法をテラーメイドする (プログラム評価論の確立)

- TruthとUseの統合

- 科学的・実用重視

III. 2000's- : 「社会的なインパクト」の要請への対応

- 社会的責任投資、インパクト投資、ESG投資、SDG s 等
- 外からの要請に対するアカウンタビリティの確保

- “impact” driven

- 外在的価値を探る

©2024 Yuriko Minamoto, All rights reserved.

27

引用・参考文献:

- 今田克司 (2004) 「社会的インパクト評価に関する現状整理」、日本評価学会社会的インパクト評価分科会
- 中村美重 (2023) 「文化事業におけるプログラム評価の課題」日本評価学会第24回全国大会発表要旨集
- 林寿和 (2024) 「IMM (インパクト測定・管理) とは何か? 実施主体は企業が、それとも投資家か?」 JANPIA 出資事業部 <https://investment-note.janpia.or.jp/n/nbec47385ac83> (2024.10.7閲覧)
- 源由理子編著 (2016) 『参加型評価～改善と変革のための評価の実践』 見洋書房。
- 源由理子・大島蔵共編著、山谷清志監修 (2020) 『プログラム評価ハンドブック～社会課題解決に向けた評価方法の基礎・応用』 見洋書房。
- 吉澤剛 (2013) 「責任ある研究・イノベーション-ELSLを超えて」『研究 技術 計画』28(1)、pp106-122
- Crossick C. & Kaszynska, P. (2016), Understanding the value of ARTS and CULTURE: The AHRC Cultural Value Project, Arts and Humanities Research Council (=2022, 中村美重訳『芸術文化の価値とは何か—個人や社会にもたらす変化とその評価』水曜社)
- Epstein, M.C. & Yuthas, K. (2014), *Measuring and Improving Social Impact*, Berrett-Koehler Publishers(=2015, 鶴尾雅隆・鴨崎貴泰監訳『社会的インパクトとは何か—社会変革のための投資・評価・事業戦略ガイド』 英治出版
- Muller, J.Z. (2018) *The Tyranny of Metrics*, Princeton University Press (=2019, 松本裕訳『測りすぎ: なぜパフォーマンス評価は失敗するのか?』みすず書房)
- Rossi, P.H., Freeman, H. E., and Lipsey, M. W. (2004), *Evaluation: a systematic approach*, 7th ed., Sage (=2005, 大島蔵他訳『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)
- Scriven, M. (1991) *Evaluation Thesaurus*, 4th ed., Sage Publications



28