

FASID 第 257 回 BBL セミナー（要旨）

テーマ： SDGs の「達成度を測る」を再考する：ポスト SDGs の「価値を引き出す評価」
に向けて

日時： 2023 年 11 月 29 日（水）12:30～14:00

場所： Zoom ミーティング

講師： 米原あき／東洋大学 社会学部社会学科 教授

出席者： 25 名（FASID 内参加者含む）

1. 発表要旨（米原教授）

- 評価とは価値を引き出すことである（extract value）。評価は、そのもの自体のよさ（内在的価値、物事の本質）、客観的に見たそのもののよさ（外在的価値、値打ち）、社会にもたらすよさ（社会的価値、意義）、の 3 つの側面から考えられる。外在的価値が重視されがちであるが、これらの視点を用いることが重要である。
- プログラムの実施と評価は表裏一体である。実施が難しいものは評価も難しく、「SDGs 評価は難しく」なっている。SDGs の取組みの難しさは、MDGs からの変化として、ゴールの本質が変化したことによる。例えば、教育分野では、MDGs のゴール 2 は「初等教育の完全普及」であり、学校教育制度の枠内にあったが、SDGs のゴール 4.7 では、市民教育や生涯教育まで含まれており、学校教育の枠を超えたものになっている。MDGs では、量的・現実的目標が設定され、指標は就学率、修了率、識字率等、測定可能なものだった（measurability）。SDGs では、理念的・理想的目標が設定され、広い目標に対して評価可能性（evaluability）が重要視されるようになった。ここでの本質的な違いとして、SDGs 以降は、抽象的な目標の unpack の必要性が生じている。
- 評価には 2 つの形態がある（形成的評価、総括的評価）。形成的評価の事例の 1 つ目として、タイの D-HOPE プロジェクトを紹介する。同プロジェクトではエンパワメント評価と Value Statement Survey を合わせた評価手法を採用した。プロジェクトの活動を実施する中でワークショップを 4 回実施し、地元の森のツアーガイドができるなどの「特技」を持つ個人の特定・経験の共有、体験プログラムのデザイン・検証、カタログ作成・配布、振り返りを行った。最後の振り返りがエンパワメント評価として行われ、評価活動を通じて参加者たちが「改善」に向けての具体策を自ら見つけ出し、構築していった。これと平行して Value Statement Survey として、ローカルの人々が考える価値の特定、ボトムアップでの指標作成、質問票作成、ベースライン調査、エンドライン調査を行った。エンドラインでどれくらいの変化がみられるかを統計分析すると同時に、調査データの因子分析を行った結果、「生活の価値」の因子として、ある地域では 7 つの要素が抽出された。これは地域によって異なるもので、地域ごとに異なる価値の多様性を

示している。

- もう一つの事例として、横浜市の ESD (Education for Sustainable Development : 持続可能な開発のための教育) スクールマネジメントでは、協働型プログラム評価を実施している。どんな取り組みも ESD に繋がると言えるが、その中で、学校が実現したい ESD とは何か、その取り組みを支援する仕組みとは何かを評価で明らかにしている。プログラム評価でいわれる「体系的な調査」とは、プログラムのニーズ評価、セオリー評価、プロセス評価、アウトカム／インパクト評価、コスト・効率性の評価、といった視点で見ることである。本プログラムでは、評価を PDCA で実施している。まずは、ニーズ評価とセオリー評価、その後、ベースライン調査を行い、中間時点でプロセス評価を行う。期末にはアウトカム評価を行い、これが来年度の改善に向けたセオリー評価となる。
- ポスト SDGs の評価に向けた 4 つの提言としては、まず、「する／される評価」から「一緒に創る評価へ (共創型評価)」という点である。サービス提供者側のパートナーシップだけでなく、提供者と利用者の共創が必要になってくる。二つめは、エビデンス概念の拡張である (間主観的エビデンス)。評価のエビデンスとして、科学的なものの対極に主観的な意見があるとしたら、その中間に、合意に基づくエビデンスや、文脈を共有する私たちにとってのエビデンス、というものも認められてもよいのではないかと思う。三つめは、「質 (構築主義) VS 量 (実証主義)」二項対立からの脱却である。質 (帰納的推論) と量 (演繹的推論) ではなく、帰納と演繹の往還により、真 (またはおそらく真であるもの) への近接アプローチ (仮説的推論) がある。RCT (Randomized Controlled Trial : ランダム化比較試験) などの方法だけでは拾い上げられないものに注目する必要があるが、そういうものをケーススタディで拾いつつ、両者のアプローチを往還することが求められる。例えば、横浜市の ESD のプログラムの指標はとても質的なものであるが、そのデータを統計的に処理し、実証的に有意性を検証している。その結果をまた質的に解釈して、来年度の指標に参考にしてしている。四つめは、評価的思考であり、先の 3 つの提言をベースにした行動への志向性である。

2. 質疑応答

- Q1. 協働型評価の課題は、当事者の負担 (時間・労力) が大きくなることかと思うが、教員のオーバーワークが問題になっているところ、その点、みなとみらい本町小学校ではどのように対処されているか。他の学校でも同様の取り組みが可能か。
- A1. 非常に大きな問題である。先生方の追加業務にならないように、教員が実施する重点研究のテーマにしてもらい、学校の教員研修の枠の中で実施したことで負担とはならなかった。同じようにできれば他校でもできるのではないかと思う。
- Q2. SDGs 全体の国連での評価方法、評価、それに対する米原先生のコメント (課題も含めて) をお伺

いしたい。

- A2. SDGs の 230 以上の指標のうち、従来の方法で収集できるものもあり、それは継続することで問題ない。指摘できるとすれば、ボトムアップの評価、ローカルレベルでは比較可能ではない評価も認められてもよいのではないかと、という点である。そうせざるを得ないゴールもある。ユニバーサルな指標で測定しようとせず、グローバルに測れない目標についてはボトムアップの評価を部分的に導入してもよいのではないかと思う。それが SDGs の本質にも整合するのではないかと思う。
- Q3. タイのプロジェクトで Value Statement Survey を開発するにあたり、開発プロセスや手法でこだわったことや悩ましかったことがあれば共有頂きたい。
- A3. 今でも難しいと思うのは、Value Statement を抽出すること、ワークショップで現場の人々からそれを拾い上げることである。文字の読み書きができない人から、Value Statement をどれくらい引き出せるのかといった課題に対しては、タイ人のローカルスタッフに助けられた。しかし、彼らもタイでは、“高い”立場の人たちであり、どこまで現場にコミットできるのかという難しさもある。ユニバーサルな指標が設定されているとプロジェクト実施者は悩まないだろう。他方、指標をボトムアップで作る場合、その指標は誰の価値を表しているのか、表し切れているのか、考える必要がある。
- Q4. 評価の結果が SDGs の当該ターゲットに貢献しているといえるのであれば、その貢献の因果関係はどのように証明できるか。比較可能なデータでは見れない、ということはあるか。JICA でも個別の事業を実施しているが、それらがどう SDGs に貢献できるか、粒度は異なるが、SDGs への貢献をどのように可視化できるかが課題となっている。
- A4. 横浜市の ESD プログラムの例は、ESD の推進につながっている（ゴール 4）。SDGs の指標で extent とあるが、実際は各国の教育省がいくつかの質問に回答し、それがポイント化されている。学校の ESD は微々たる貢献で遠い距離ではあるが、4.7 の指標の一部に貢献しているといえる。タイのプロジェクトでは、SDGs のつながりでいえば、働きがい・経済成長へ貢献する取組みだと思う。このプログラム終了後、タイ関係者がタイの内務省から participatory governance という賞を受賞したことからも、SDGs のゴール 16、17 への貢献といえるかもしれない。それでも直接的な影響を与えたとはいえず、インパクトは不明瞭である。個々の評価がビッグピクチャーの中でどのように位置づけられるか、ロジックでつなげる努力をしなければならない。230 以上の指標はマクロデータの指標が中心であるが、そこにインパクトを与えるかどうかということより、voluntary review の方で汲取っていくのもよいと思う。
- Q5. 形成評価、アジャイルガバナンスとも重なるものではないかとも思う。タイの事例に関して、

KPI (Key Performance Indicator : 重要業績評価指標) と現場との乖離をなくす努力をしたとのことであるが、日本でも同様の課題がある。タイでは KPI につながるかためにどのような努力をしたか。

A5. ボトムアップの指標は、KPI に直接関連する情報にはなっていない。所得、売上げという指標も、データ自体が帳簿で管理されていなかったり、そもそも定価がなかったり、計算が合わない。KPI で現場の状況を把握するのは無理である、ということを伝えるのも意図していた。求めること自体に自省を促す、違うアプローチも必要、というメッセージとして伝えた。前述のとおり、participatory governance の受賞は、それでもよいと理解されたのではないかと思い、ありがたかった。

Q6. エビデンスの概念拡張について、もっといろいろな業界や分野で議論が進んだらよいと思う。間主観的なエビデンスを、その文脈に関わっていない第三者に共有し、理解してもらうために、どんなことが必要か。これまで関わられたプロジェクトなどで、評価結果のコミュニケーションの工夫があったら教えて頂きたい。

A6. 重要な点である。間主観的なエビデンスは共有できる範囲のものだが、範囲外にも説明が必要になる。例えば、可能な限り、量化できるところは量化する、第三者にとってわかり易くて興味を持ってもらえる報告とするという工夫がある。横浜市の ESD のプログラムでも、一義的には先生方自身が取組んだ参加型の評価であるが、文科省や市への報告にあたり、データによる可視化、質的データのテキストマイニング等、分析結果が一目でわかり易くなるよう準備した。発信の仕方を工夫して、範囲外の人に伝わるようにする。認識、言葉が違う人に評価結果をどう伝えるか、意識していく必要がある。

Q7. 形成的評価と共創の評価について、PCM¹ (Project Cycle Management) が「本来」めざしているモニタリング・評価との基本的考え方との共通性を感じさせられた。タイのプロジェクトの Value Statement Survey での 7 つの分類は、普遍的な分類として何かに準拠したものか。あるいは、現場の状況をもとに開発されたものか。分類の仕方によって評価結果が変わってくるようにも思えるが、どのようにこの分類方法を策定されたのか教えて頂きたい。

A7. 本来は、本日発表した内容は PCM が目指すものと不整合性を持つものではない。発表の 7 分類は、因子分析の結果である。20 の value statement があり、それが事前事後でどのように変わったかを分析した後、その 20 の変数を因子分析した結果である。チョンブリ県では 7 因子であったが、他県では違う数となる可能性もある。因子分析の特徴である。統計学では、分析者が仮説を作り、そのモデルを検証するという仮説検証型の (confirmatory) 分析方法が一般的だが、因

¹ PCM 手法とは、開発援助プロジェクトの計画・実施・評価という一連のサイクルを「プロジェクト・デザイン・マトリックス (PDM)」と呼ばれるプロジェクト概要表を用いて管理運営する方法のこと。

子分析は探索型 (exploratory) の分析である。「データに語ってもらう (Let the data talk)」分析手法である。

以上