

評価ファシリテーションガイドの開発に向けた試行的研究

清水 潤子（武蔵野大学 人間科学部 講師）

要約

本稿は、多様な主体による事業活動の評価を推進していくための行為として、「評価ファシリテーション」に着目する。そして、事業活動の関係者たちが自らプログラムを改善・変革していくための評価ファシリテーションガイドの開発・実装について報告するものである。本稿は三部構成となっており、第一に、ガイドの開発にあたって行った評価や評価ファシリテーションに関する先行研究をもとにした基本概念や理論について整理する。第二に、先行研究等に照らして作成した評価ファシリテーションガイドのプロトタイプを実装した研修内容について概説する。第三に、同研修に参加した参加者と企画者へフォーカスグループインタビューを実施し、その分析を通じてガイドを今後ブラッシュアップしていく際のポイントを提示する。

1. 研究背景

複雑な社会課題の解決にあたっては、一個人や一組織、一つのセクターで生み出せる成果には限りがあり、現代社会においては様々な環境において協働や連携の必要性が謳われている。このような現場において、ともに目指す成果や価値創出、目標の達成のために、事業活動の改善や学習を目的とした協働的な評価実践の定着・促進が必要である一方、「評価」自体への苦手意識から、それがなかなか進まないという課題がある。

このような課題に対して、事業活動に関わる多様な主体が評価対象となる事業活動の在り方や評価軸について、ともに検討していくことが重要であり、関係者とともにワークショップ等をデザインし、対話や問題意識の共有等を通じて自らの事業を評価していくことが期待される。

本研究では、このような評価を推進していくための行為として、「評価ファシリテーション」に着目し、事業活動の関係者たちが自らプログラムを改善・変革できるようになるための評価ファシリテーションガイドの開発と実装を

行った。本稿は、このガイドの開発と実装のプロセスについて報告するものであり、第一に、ガイドの開発にあたって基盤となる評価や評価ファシリテーションに関する基本概念や理論についての先行研究の確認・整理を試みる。第二に、概念整理等に照らして作成した評価ファシリテーションガイドのプロトタイプを実装した研修内容について概説する。第三に、同研修に参加した参加者と企画者にフォーカスグループインタビューを実施し、その分析を通じてガイドをブラッシュアップしていく際のポイントを提示する。

2. 評価に関する基本概念の整理

本稿が着目するのは「評価ファシリテーション」であるが、それがどのようなものであるかを理解するためには、そもそも評価や評価ファシリテーションとは何かについての理解を深める必要がある。本章ではそれぞれの定義や概念を確認する。

(1) 評価とは何か

評価の定義については多くの議論があるが、評価の英語表記が“Evaluation”であり、その原義はラテン語の *ex-*（外へ）と *value*（価値）を組み合わせた言葉である。そこには、評価対象の価値を引き出す、見出すという意味が含まれている。学術的な領域における定義として代表的なものは、「評価は、物事のメリット、值打ち、意義を体系的に明らかにすることである」(Scriven, 1991, p.39) という評価学の大家であるスクリヴェンの示した定義がある。同定義は全米評価協会でも会員や評価実践者、研究者による議論の末、評価を定義する文言として用いられており、一定学術的・実践的にコンセンサスを得ている定義だと考えられる。

「体系的に明らかにする」という点でリサーチ（調査・研究）と混同されることも多い評価であるが、実質的な評価は「何が起きているのか、○○とは何か？」ということを問うような記述的な研究よりも一段階難しいもので、このような記述的な問い合わせに加えて、それらはどの程度良いものなのか？それらの成果は十分なものなのか？というような価値判断を伴う行為である (Davidson, 2011)。

また Schwandt (2015) は、狭義では「評価とは、プログラムや政策がどのように機能しているか、そしてそれが効果的であるかどうかを判断するために、研究手法を活用して情報を収集するプロセス」と定義し、広義では「評価に特化した手法を用いて何かの価値を判断する批判的思考の一形態」だと定義する。社会科学の実践・研究領域において、イメージされたり議論されたりする多くの評価は Schwandt (2015) の狭義の評価の意味に近く、社会的な評価の活用イメージもここにあると言えるだろう。しかし、評価の理論家がそれ以外の定義をする背景には、スクリヴェンが評価は「価値づけの科学」であると言うように (Shadish, Cook, & Leviton, 1991)、評価対象の善し悪しを判断するという価値判断の在り様についての学問であり、評価基準や価値軸をどのように設定し、それらはどうあるべきなのかに関する議論に本質的に力点を置くからだと考える。

藤川 (1994, p.39-41) は、「価値」は個人的価値〔善〕、社会的価値〔正〕、自然的価値〔真〕の3つに大別され、これらの善(悪)、正(邪)、真(偽)を論じる際には、それがいかなる決定基準に準拠しているかを確認し、その理非をただすことが先決であるという。このプロセスがない議論はナンセンスであり、合理的な解は得られず、すべての価値判断は選ばれた決定基準に相対的なものであり、ゆえに判断の内容よりも、判断の価値決定基準が何かを問うことが必要である (藤川, 1994)。見田 (1966) は、藤川 (1994) よりも、より詳細な価値の類型を示しており、価値の諸類型は伝統的な「真・善・美」という図式のように相互に無関係な概念として単純に並列されるべきではないと指摘した上で、自己の今と未来、社会の今と未来というように時間・社会的な側面で四象限に分類でき、それぞれを「快」価値 (↔苦)、「利」価値 (↔害)、「愛」価値 (↔憎)、「正」価値 (↔邪) としている。また Graeber (2001/2022) は、価値に関する学術研究のルーツやその言葉の意味は、社会学、経済学、言語学、人類学という領域によって異なることも指摘している。このようなことからも、Schwandt (2015) が、評価を批判的思考の一形態であるとする背景には、この価値基準の多様性を考慮した上で、そのような視点を評価という行為が本来もたらすべきものであることが読み取れる。

また、この評価における批判的思考を評価学領域では一般的に評価的思考

(evaluative thinking) と呼び、このような思考を含む評価を行う者の評価能力 (evaluation capacity) を高めていくことの必要性を主張する代表的な評価理論家としてパットンがいる。Patton (2008) は特に評価結果の使い手を意識し、使い手が異なれば評価のニーズも目的も変わり、それによって何を問うか、アプローチをどうするか、価値判断の基準をどうするかということは多様であることを示している。評価を文脈やそこに集う人々や彼らのニーズに依存するものとして捉えていることが特徴として伺えるが、ゆえにそのような中でいかに価値の基準を持ち、どう判断していくのか、また現在想定されているような価値軸を疑えるような思考をもつことが重要であるという点で、Schwandt (2015) の観点とも重なる。

これらの評価理論家たちが示してきた代表的な評価の定義について Podems (2019) は、それぞれにスタンスは異なるが互いに排他的ではなく、評価という試みが体系だった形で行われる必要があること、何らかの価値を判断することの 2 点を挙げ、また結果的に政治的であるということを通底する共通点として指摘している。

(2) 評価者とは誰か

では前項のように定義される評価は、誰によって行われているのか。評価という営為は、私たちの日々の生活のあらゆるところに存在している。身近なところでは成績評価や人事評価、福祉サービスの第三者評価というように、「評価」という名前がつくものから、オンラインストアにおける購入者の口コミやレビュー、レストランの格付け、病院ランキング、何らかのコンテスト、SNS のいいね！やスキ！、芸術鑑賞、スーパーの販促コーナーでの味見など、「評価」とまでは言われないものの、何らかの価値基準に沿って（それが主観的・客観的であるかに関わらず）対象の善し悪しを判断する行為まで様々である。公共的な領域で行われる政策評価や行政評価、大学や教育の評価、ODA の評価のように、多額の予算や多くの利害関係者が関わる政策・事業活動に対して評価を行う場合もあれば、市場や私的なニーズを満たすための評価も存在し、「評価」は社会においてそれぞれの文脈で多義的に認識されている。本稿においては社会課題の解決や社会的価値を創造していくための評価の在り様を探索していく

という背景から、本項では評価の学術研究が進んだ公共的・社会的な領域に限定して考えてみたい。

Schwandt (2015) は、評価者や評価専門家といわれている人々の多様性を指摘したうえで、3つの類型を示している。第一に調査機関に雇われたり、大学において評価実践をある程度行ったりしている経験のある評価者、第二に初めてプログラム評価の領域に足を踏み入れて知識やスキルの向上を目指す初心者、そして第三にトレーニングを受けていない人が、自分の職責の一部として評価の実施を任せられ、その仕事をするために必要なことをやろうとしている人 = accidental evaluators である。

①米国での状況

評価者の伝統的な役割とは、第三者的で客観的な調査を行うというもので、プログラム等のインプットやプロセス、アウトカムに関する的確な評価報告書を作成し、事実の理解を追究する者だと認識されていた (Weiss, 1998/2014)。特に調査が大規模で複数の調査対象地が必要である場合は、現在でもこうした評価者の役割が中心的な位置を占めており、より高位の当局によってプログラムの継続性を判断するような評価において、こうした傾向が見られる (Weiss, 1998/2014)。このような評価者においては、評価者が絶対的な中立を保つことは不可能である (Schwandt 2015; Weiss, 1998/2014) が、良いデータ、健全な分析、率直な報告といった科学的研究が重点を置く慣習に従う傾向がある。

しかし、評価者の正しい役割とは何かということについて異なった意見も出現しており、エンパワーメント、公正、解放といった概念の出現に呼応して、批判的な友人、協同調査者、ファシリテーター、問題解決者という評価者の役割についても議論されるようになった (Weiss, 1998/2014)。このような背景には、評価が評価知識の構築に力点を置いていた時代から、評価知識の活用に重きを置いた理論が登場し、実験デザイン手法に限らない定性手法を組み込んだ評価理論の構築が進んだことが認められる (源, 2016)。社会科学の量・質論争、実証主義・ポスト実証主義に対する社会構築主義・実利主義のパラダイムやポストモダニズムの考え方に対する評価アプローチが出現することで、評価者の役割やスタンスも多元化していった。

Podems (2019, p.215) は現代的な評価者の役割をリスト化しており、その中に描かれる役割には、interpreter; advocate, change agent; anti-bully; critical friend, collaborator; coach; cruciverbalist; developer; diplomat; learner; logician; manager; scholar; lecturer/professor; detective, sleuth; describer; designer; reporter; novelist, writer; researcher, social scientist, methodologist; judge, magistrate という 19 のジャンルがある。前述の伝統的な評価者から想起される役割としては researcher, logician, describer, judge, reporter などがあり、これらは従来の評価を行う際に評価者に認められる重要な役割であるが、評価の発展に伴い、役割の拡張がみられる。

全米評価協会のタスクフォースでも、評価の発展に伴う評価者役割の多元化に着目し、評価者とはどうあるべきかについてや、評価者にどのようなコンピテンシーが求められるのかについての議論が進められてきた。5,000 人を超える会員 (American Evaluation Association, n.d.) によって成り立つ全米評価学会では、1994 年に最初の評価者の行動規範 (guiding principles) が定められて以降、5 年に 1 回見直しが行われる中で、社会の要請にともなって評価者の行動規範も発展してきた歴史がある。また、2015 年から 2018 年の間にこれらのメンバーを対象としたアンケート調査や、年次大会等での会員との議論を経て、現在 5 つのコンピテンシー領域 (professional practice; methodologies; context; planning and management; and interpersonal) が定められている。紙幅の関係で、同行動規範の進化のプロセスやコンピテンシー領域が制定された背景を詳しく述べることはできないが、元来、頑健な科学性に則り、評価者が中立、独立の立場を強く示していた時代における評価者像の大きな違いとして、より評価を必要とする人や、利害関係者とのコミュニケーションが奨励されていることがある。また評価のプロセスに多様な主体が関わるようになる中で、利害関係者とともに価値判断を行い、評価を必要とする人々や組織の評価力を高めていくという社会的要請から、評価者の役割が拡がっていたことが伺える。Podems (2019) も自らが評価者として意義深いと感じる役割として、教育者、ファシリテーター、ネゴシエーターの役割を挙げ、これらはよく評価の依頼者や参加者、利害関係者などとのかかわりの中で話題となるが、同時に評価チームをリードしたり、評価チームの中でアクティブに活動したりする上で不可欠

な役割だと指摘する。

②日本の状況

米国とは異なり、日本は一部を除き、組織内部の評価担当部門の職員自身が推進する内部評価が主流であり、予算や人員など諸般の事情から、施策や事業を実施した職員自身による自己点検評価である場合が多い。政策評価や行政評価の歴史も北米に比べると浅く、未だ実施していない自治体もある（池田，2021）。行政と同じく、米国で評価が発展した民間非営利組織の支援や助成の分野においても、助成事業の成果や業績を評価するために、監査が入ることはあっても、第三者評価・外部評価の目的で評価が入ることや、その費用が認められている事業は少ない。一方で昨今の、エビデンスに基づく政策立案の流れや、公共領域における成果志向が、社会的な評価ニーズを増していることは明らかである。

一方、日本は、世界で見ても評価者の認定制度を持つ数少ない国の一いつであり¹⁾、評価に関する知識を持つ人は未だ少ないものの存在している。しかし、一定数専門的な知識を持つ研究者や調査者が、事実特定のための調査（規模の大小や質は問わない）を行っていたとしても、評価が自己点検や内部評価によつて行われることが多かったり、特に評価の発展が諸外国で認められる政策や行政の領域においては、職員も都度異動等をしたりするため、評価の実施もSchwandt（2015）がいう *accidental* にならざるを得ない状況がある。

しかし、この *accidental* な形でも評価に直面した人々が、評価の本質的な意味を理解し、自身や利害関係者が関わる活動の評価を体系的に促進し、その結果を活用することができれば、公共的な活動をする人々が、より事業活動を良くしていくためのパワーを持つことができるのではないか。

（3）小括

ここまで評価の定義に関することと、その評価者の在り様について述べてきた。評価は価値判断に関する科学であり、いとなみであること、そしてそれを判断する主体としては専門職から経験の浅い、もしくは突然そのタスクが降ってきた人によって當まれており、またその方法については、評価の目的やアプ

ローチの多元化等によって様々であることがわかった。公共的な領域において私たちが直面する課題は複雑であり、そのような課題にたったひとりの専門職や組織が立ち向かうことでできることは限られている。また、仮に立ち向かったことや達成できたことを評価する際に、日本においてはまだまだ外部・内部評価のリソースは限定的であり、それを行う文化も未成熟である。しかし、世の中のプロジェクトや事業活動の数だけ、明示的もしくは潜在的な accidental evaluator がいるとしたら、その人たちがより評価活動に意義を見出し、評価の力を借りて事業活動を良くしていく可能性はある。事業活動には様々な利害関係者がいるとしたときに、そのような利害関係者とともに、評価をしていくことはできないか。

3. 利害関係者とともにを行う協働・参加型評価とファシリテーションについての概念整理

(1) 協働・参加型評価

利害関係者が評価活動に参加したり、協働したりしながら行う評価についての研究や実践報告はこれまで数多く蓄積されてきている (Cousins & Chouinard, 2012; Cousins et al., 2020; 源, 2016; Wingate et al., 2025)。多様な評価理論家や実践家によって提唱してきた利害関係者の協働や参加を伴う評価理論やアプローチは「評価における協働型アプローチ」(Collaborative Approach to Evaluation: [CAE]) (Cousins et al., 2020, p.6) としてまとめられ、「評価者と非評価者が協力して評価的知識を生み出すことを基本要件とする評価アプローチ」として今日では包括的に整理されている。

前章で示した評価の定義等に照らして考えると、公共的な利益を求めて事業活動を開拓する中においては社会的価値(正)が求められることになる。しかし、公共や社会といつてもその担い手は多様であり、例えば多セクター連携の事業で、行政・企業・非営利組織が協働するとなると、それぞれのセクターにおいて規範となる価値軸が異なるということが当然起こる。また、ひとつの組織において行われているサービスや事業であっても、そのサービスの受け手や提供側の現場スタッフや上司、他の利害関係者で異なる価値軸やパースペクティブが存在し、このような異なる主体が関わる事業活動を評価するために、

どのような評価軸が必要なのかを議論する必要がある。

評価学の発展のプロセスにおいて、「協働」や「参加」の概念が重視されたのは、評価の目的に照らして「誰が評価者（＝価値判断をする主体）たるべきか」ということに関連している。独立型評価と特徴を異にするものとして価値の捉え方（valuing）の在り方があり、利害関係者が評価プロセスに参加する参加型評価では、関係者間の協働関係から創出される価値を重視し、評価設計から収集したデータの分析に至る一連の流れを利害関係者が共に行うことで、利害関係者の認識と懸念に応答した評価を行うことになるとされている（源，2016）。

日本においても、このような観点から行政や非営利セクター等における政策やプログラムにおいて協働や参加の概念を重視する評価実践が増えており、プログラム評価や Most Significant Change (MSC) という評価アプローチやテクニックを用いた評価において行われ、そこには関係者間の協働関係から価値を創出できるように、評価者ないしは評価専門家がファシリテーターの役割を積極的に担い、評価プロセスを推進するというものである（源，2016; 田中，2010）。

（2）評価ファシリテーション

評価ファシリテーションについて Patton (2018, p.2-3) は、ファシリテーションの広い分野における専門的な分野で、プログラム評価に携わる利害関係者との協働における特有の課題に対応するため、一般的なファシリテーションの知識と技術を適用・適応するものだとし、このようなファシリテーションの目的を、評価の関連性、信頼性、意味合い、および有用性を高めること指摘している。

評価と同様に、ファシリテーションについてもそれが実施される分野は多岐に及ぶ。まちづくりや企業・団体の組織開発に留まらず、日々の会議や教育現場はじめとする多様な現場において、ワークショップやファシリテーションに注目が集まる今日、ファシリテーションについて学ぶ書籍や研修機会なども増えている。それぞれにおいて多様なファシリテーションの定義がなされているが、小田（2022）は、ファシリテーションの簡潔な定義を「グループによる知的相互作用を促進する働き」としている。また、小田（2022）は、討論や会議といった場で司会進行をする人が場を仕切ったり、論点を提示したり、結論をまとめたりするような行為は「モデレーション」と呼ばれる方法であり、議論

の内容に焦点を当て、論理的な結論を引き出すことを目的とするものに向いていいると指摘する一方、ファシリテーションは、議論の内容だけでなく、その場に集まる人々の間で起こるグループ内の「プロセス」や文脈、関係性、場の質に焦点を当てて、メンバーたちの新たな理解や関係性を生み出していくことを目指すものだと指摘する。このような点から、多くの利害関係者が集まり会議や討議をするような場では司会が求められ、何かしら一つの答えを出していくことが目的とされやすい。しかし、評価「ファシリテーション」であり、評価「モディレーション」ではないということを考慮すると、評価ファシリテーターに求められることは、単にプログラム評価に関する知見を教師のように教えたり、アドバイスを提示したりするためだけに関わるのではなく、その場に集まる人々と、「価値づけ」のプロセスに身を置き、それぞれの発言の意図やメンバーのダイナミクスなどを考えながら、メンバー同士のやりとりから学べるようにすることが必要だと読み取れる。

小田（2022）が示すような知的相互作用、ないしは人々の相互作用によって人々の行動や思考が影響を受けるという考え方の背景にある社会的相互依存理論（social interdependent theory）は、評価学やファシリテーションに関する学術領域の根底にある概念として共通している（Fierro et al., 2016）。また評価のプロセス活用や組織学習の理論等も評価という行為とファシリテーションという行為をつなぐものとして指摘され、評価とファシリテーションは、振り返りを促し、学習を促進するという共通の目的を持っているとされる（Fierro et al., 2016）。

4. プログラム評価に評価ファシリテーションを導入する

これまで述べた先行研究や、協働・参加型評価に評価者ないしは評価専門家として関わってきた経験を持つ研究者の実践的な知見を集約し、2023年からプログラム評価²⁾の特に形成的評価のフェーズに特化し、福祉的ニーズを持つ人々を対象にする事業活動を展開している事業者向けに、評価ファシリテーションの概念やノウハウについてのガイドを取りまとめるプロジェクトを開始した。筆者に加え、2人の共同研究者と内容を構成、執筆した。2023年6月18日（於：明治大学）にガイドのプロトタイプを作成したものを、公募によって集まった

評価ファシリテーション研修の参加者に提示し、そのガイドをもとに評価ファシリテーションを学ぶワークショップの機会を持った。

(1) ワークショップの参加者

ワークショップの参加者は総勢 16 名であり、プログラム評価に関しては、本研修の前に行ったプログラム評価の基礎に関する研修を全員が受講済みであり、プログラム評価やロジックモデルの基礎的な知識を持っているメンバーであった。参加者の属性は多様であり、コンサルタントとして評価支援を行う立場の参加者や、福祉系団体や NPO 等で福祉ニーズを持つ人々を対象に事業を行う実践者や同組織の管理職、研究者等も参加した。

(2) ワークショップの内容

参加者には事前に事例を提供し、それを読んでもらってからワークショップに参加いただいた。ワークショップでは実在する福祉サービスプログラムの事例を加工し、事業を取り巻く環境や内容、目的、課題や、その事業を提供する組織の状況等の情報についての内容を盛り込んだ。

ワークショップ当日は、参加者が 1 グループ 5 – 6 名で構成される 3 つのグループに分かれ、プログラム評価の形成的評価の中でも重要なセオリー評価の一部として、事例に描かれている情報をもとにロジックモデルを作成する場面を想定し、各参加者が「最終アウトカム」「中間アウトカム」「直接アウトカム」「活動」について議論する場のファシリテーションを行った。



写真1:ワークショップの様子

筆者と共同研究者は、同ワークショップの企画者としてモデレーターとプログラム評価およびロジックモデルについての簡易的なレクチャーを行い、各グループによるワークショップの実施中は、各グループの様子を観察し、ファシリテーターとしての振る舞いや場の様子を観察し、気づいたことなどをファシリテーター役にフィードバックするという役割を担った。

5. フォーカスグループインタビューの実施

(1) 概要

ワークショップの実施後に、実際ワークショップで学んだ評価ファシリテーションを現場等で使う機会があったか、また今後ガイドやワークショップについて改善する必要があれば、どのようなことがあるかについて参加者3名と企画者3名による半構造化フォーカスグループインタビューを行った。参加者のリクルーティングについては、2023年6月18日の研修参加者全員にメールにてフォーカスグループインタビューへの参加協力の依頼を行い、協力を希望頂いた3名を採用する形をとった。

フォーカスグループインタビューを研究法として用いた背景としては、人々の経験や知識を明らかにしようとするときに特に有用な方法であり、人々が何を考えているかやなぜそのように考えるのかについて明らかにできるということや、経験や考え方の似た人々が集まるため、参加者にとって話しやすい環境となることが挙げられている (Liamputtong, 2020/2022)。具体的な質問項目としては、「ロジックモデルづくりのファシリテーションを実際に体験してみて、その実践に必要なスキルはどのようなものだと思ったか」、「ロジックモデルづくりのファシリテーションをやってみて（やろうとして）難しかったところ（難しいと思うところ）はどのような点か」、「ロジックモデルづくりのファシリテーションと日常的なファシリテーションを比較しての共通点や違いを感じるとしたらどのような点があるか」の3点を提示し、相互の創発的なディスカッションを重視した。フォーカスグループディスカッションの実施にあたっては Liamputtong (2020/2022, p.81-107) を参照した。本研究の対象者には、研究目的、方法、個人情報の保護について、文書と口頭で説明を行い、同意を得て、2024年1月29日にオンラインで実施した。参加者について表1

にまとめる。また、フォーカスグループディスカッションの進行は、筆者が行った。なお結果の記述に関しては、参加者が特定されることを防ぐため、一部抽象化して表記する。

表1:フォーカスグループディスカッションの参加者

カテゴリ	ID	主な所属	性別
参加者	A	社会福祉法人	男
参加者	B	特定非営利活動法人	男
参加者	C	大学	男
企画者	D	大学	女
企画者	E	大学	女
企画者	F	大学	男

(2) 結果と分析

インタビューの参加者の内、研修に参加した参加者からは、評価（ロジックモデルづくり）についての知識やスキルについてとファシリテーションの技術が必要であることが述べられたが、中でもファシリテーションを行ってグループダイナミクスに対峙しながらロジックモデルづくりを完成させなければいけないという課題を認識し、難しさを抱えているように見受けられた。具体的には、以下の通りである。

A：参加者が、何ていうのかな、フラットな立場であったりとか、あるいはファシリテーターの心理的な安全性が確保されているような場ならいいんですけど、例えば、ワークショップの仲間というか、チームの中にちょっと発言が強い人がいたりとか、この人ちょっととな・・・みたいな、部長、上席がいたりとなったときに、ちゃんとできるかなってちょっと非常に不安かなって思います・・・(中略)・・・コンサルとか、ある程度身分がある人って言ったらあれだけど、例えば大学教員とか、そういう人じゃないとなんか、ぐっと進めて行くのは非常に難しいのかな、自分でできるかなっていうのは、正直感じた所です。

B：アウトカムを決めていくときに、その参加者の人たちによっては持つ

ている情報量が違ったり、それぞれの立場がある中での話し合ひっていうところが、なかなか立場をはがしきれない部分があって・・・そうすると、こう思っているんだけど言えないみたいな・・・(中略)・・・例えば経営層の人には、経営層なりの情報があったり、現場レベルの方には現場レベルの方の情報量があると思うので、その違いがある中で、一つの目標を決めていくっていうことは、結構難しいんじゃないかなって思いました。

C：今、行政と○○と◆◆の3者協議で福祉の実践を行っているんですけども、その中で、◆◆としては、この事業をどう進めていくかっていうところに割と重きを置いていると。だけど○○は、いかにそれを全国に広めるか、というか、もっと言うと、いつになつたら法制化するんだみたいな話になる。おそらくあんまりそこら辺話はしてないんですけど行政としてはそんなに大きくっていう話じゃないけど、少なくとも市民サービスとしてこういう事業っていうのを行っていくときに、どうしていこうかみたいな。

このように評価ファシリテーターとして、評価ワークショップに参加するメンバー間のダイナミクスや立場性の違いにどのように立ち向かおうかと考える中で、難しさを抱えていた。一方、このような発言に触発されて、参加者の内、研修の企画を行ったメンバーからは、以下のような発言が認められた。

D：多様な人たちとロジックモデル作りを実際やったりとかするんですけど、ロジックモデルづくりのワークショップを2回、3回って積み重ねていく中で、何か皆さんの中に納得感があるものっていうか、得られるものって、目に見えるロジックモデルが整理できたっていうよりか、何かこういう対話を通して、あの人ってそういうふうに考えていたんだとか、自分たちが今まで地域福祉計画とか総合計画とか作る中でモヤモヤしていたことを、何か議論する場ができたことに感謝されたりだとかっていうことは結構多くて。

E：みんなで、関係者で一緒にっていうのは、関係者とともに、関係を作りながら、協働でやっていく、そういうプロセスの一つとしてワークショップ。何かそのプロセスで関係作りができるっていうのをすごい期待されていると思うんですよね。で、実際にそういう結果になる。・・・(中略)・・・なので、評価を一つの道具として、あるいはその評価の思考みたいなことを、みんなで共有していく場みたいな捉え方を結構しているな、と感じました。

F：ロジックモデルを作るっていうのが単純なゴールであれば、ワークショップじゃない方法とかも結構あると思うんですよね。担当者がその、作っていくと。それを会議とかにして見ていただいてですね、いいですか悪いですかみたいな、やり方は当然あるわけで。あえてこのワークショップでロジックモデルを作るというのはそこで対話をして、みんなで合意する。一つのセオリーをしっかりと組み立てようっていうどちらかというと綺麗なロジックモデルができるってアウトプットではなくて、そのプロセスの中でいろんなものを産んでこうっていうそういうことを大切にしているんだろうなと私も思っています。

E：私も1回のワークショップで綺麗になるっていうのはまず無理だっていう体でいきます。ただそこで何か険悪なムードになったり、対立したりっていうのだけは避けたいというのあります。

研修参加者側と比較して、企画者側の発言は、評価ワークショップとして勿論ロジックモデル作りは進めるものの、参加者間の対話を創発し、またその部分にファシリテーターと時間を費やしてその場のダイナミクスを変えていき、そのプロセスに評価ワークショップの意義を見出していることが見受けられた。なぜこのような違いがでるのか。企画者側は過去に比較的グループダイナミクスが成熟した環境で評価ファシリテーションをしたからではないか、という仮説も立つ。参加者間のコミュニケーションが活発で、フラットに議論ができる環境である方が、このような議論は進みやすい可能性は高い。今回の企画

者は、短くとも8年以上の評価伴走・評価実務経験を持つ中で、それぞれが難しいと思うグループのワークショップに遭遇した経験がある。しかし、参加者との相互的なインタビューのやり取りを通じて、たとえば筆者においては、評価を必要とし、ともに評価のプロセスを進めている人々の評価の目的や意義を再確認しながら、評価ファシリテーターとしての関りの優先事項を変えたり、評価とともに進めるコアメンバーの力量や関係性のアセスメントを行い、軌道修正を図ったりしたという暗黙知の自覚につながった。

またこのような考察を通じて、評価を必要としている人と評価ファシリテーターの位置取りという観点でフォーカスグループインタビューを分析すると、同じ参加者間でもファシリテーターと組織の位置取りが異なっていることが読み取れた。例えば、

C：（自分たちの）事業に関連するファシリテーションプログラムとの比較で考えてみたんですけど、共通していることの一つはファシリテーターの専門性が高いということ。もう一つは黒子に徹する。つまりファシリテーターが答えを出すわけではない。導くというか・・・（中略）・・・で、その際に何ていうか、どの程度までなんていうかな。介入してくるっていうちょっとと言葉が違うと思うんですけども、つまり皆さんにおまかせしますよと、完全にしてしまうと話が進まないような気もしますし、だからと言ってファシリテーターが引っ張りすぎちゃってもよくないのかなっていうところで、その何ていうか、さじ加減というか交通整理っていうところで、どの程度まで何かファシリテーターが入れてるのかな、なんていうところがうん。どうなんだろうって。

というように、ファシリテーターの位置取りとして、答えは出さない、という認識がある。つまり評価でいう「価値判断」はしないということを認識しているように伺える。

Bについては、評価ファシリテーション以外の場でファシリテーションを経験することがあり、

B：結構ファシリテーターがいるグループワークとかワークショップみたいなものって、最終的な学習目標みたいなことと、それをするために取り組むワークみたいなのがちょっと違ったりしてて。何かカードゲームを使うを通じて、関係性を深めるみたいなことを目標にするときって、初めから今日のワークショップで関係性を深めましょうみたいにすると深まらなかつたりするんですけど、このロジックモデル、これを作りますっていうのを初めに、言った上で、それに向かって話し合っていくと、何か正解があるんじゃないかなって思ったりしていて。

という発言に認められるように、評価を通じて、組織の関係性を深めていきましょうということを明示してワークショップを行っても、関係性を深めることは難しいことを示している。しかし、企画者の捉え方のように、関係性を深めることがワークショップの目的ではないが、ともに関わる公共的・社会的な事業の価値の議論のプロセスを通じて、関係性を深めていくということにファシリテーションが期待されていることが感じられる。一方で、

B：ロジックモデルを作るっていうプロセスと、できたロジックモデルを活用していくっていうところと、あって、その活用していくところは、自団体で自走できるような形がとれるといいなと思ったりはする一方で、作るっていうプロセスを、自団体でやっていくっていうときに、…(中略)…なかなかそれを自団体のみでやるのは難易度が高いのかなっていう気が個人的にはしました。一つは、その評価の知識を誰かが得ていないといけないっていうことがあって、多分誰かそのファシリテーターをやる人がそういう勉強をして覚えるってなったときに、その人の役割をどういう立場の方がやるかどうかにも結果が大きく影響していく可能性があるんじゃないだろうかって想像したのが一つ。

というように、自組織で自分が行うとなったときに、その影響を懸念する声も聞かれた。同様に A からは、

A：ロジックモデルのファシリテートの場合、例えば最終アウトカムを決めるときに、複数出た、アウトカム候補から1個決めないといけないですよね。そうすると、決定するっていうところが非常に重要なと思うんです。

A：例えば、インパクトと最終アウトカムって違うのか（みたいなことを）良く理解できるかできているのかな、と・・・中略・・・（参加者から）こうですよねってツッコミ入ったときに、ちゃんと何かみんなが納得感を得られる形で説明できるんだろうかとかいうことで、ファシリテーターの能力としてその、何ていうのかな、単にその場をまわすっていうことと、その評価全般に対するその理解だったりとか、自分なりの言葉にして出していく能力っていうのも求められてるのかな、と。

という発言が認められた。特に、今回のインタビューの参加者は自分が組織に所属し、管理的業務を行う立場でもあり、自組織において自分が評価ファシリテーターとしてファシリテーションを実施したらどうなるかという思考にもとづくと、自身はファシリテーターでもあるが、当事者として決めないといけない人という役割も求められる。そのような立場を考慮したときに、ファシリテーターである当人が、組織の中で他の役割を担うと、ファシリテーターとしての自分と参加者としての自分の間にジレンマを抱えることになる難しさがあるといえる。この関係性の境界があいまいであると、ファシリテーターが「決める」という役割を担わされることにもつながりかねない。

しかし、このように組織の意思決定に活かす評価の視点の重要性に関する声も聞かれた。

B：なんかやっぱり日常的に必要とされるものになるためにはどうしたらいいかっていうこととか、何かそれを話す機会っていうのが、職場の中でどれだけ作れるのかみたいなところもあるのかなって思いました。ここはやっぱり作るプロセスに参加をしている人と、していない人みたいなところで温度差が出てきたりもすると思うので、やっぱり多くの方が参加をす

ることで、意識的にこのことを考え続けるみたいな雰囲気を作れるのかなと思います。

A：でも全然評価の知識のない職員さんと一緒にやってくっていうのは、先ほどいろいろな人を巻き込むって話もあったんですけどそれがもう非常に難しいかな・・・何でこれやんないといけないのっていう、ところからまずそこを合意を得るっていうことが、ますすごく難しいっていうか、できるかなって思います。

C：自発的に結成された集団、例えば私達がやっているその事業がまさにそうで、これをやりたいからやっているっていう場合と、例えば職場でこれをしなきゃならないっていう場合とは全然違うのかなと。前者は例えば先ほどの共通理解とかっていうところでも、目指すところっていうのは一定化しているから、例えば評価っていうのも重要だよねとか、これロジックまで作らなきゃなんないよねっていうところで割と理解が得られやすいところがある反面、仕事の一環としてやっていますっていう場合だと、場合によると、何でそんなことしなきゃなんないのとか、ちょっと何か面倒くさいよねとか、あんま興味ないよねみたいな話になる場合もあるのかな。

B：例えば地域の人も一緒に参加してもらう必要があるのであれば、声をかけなきゃいけないし、そうじゃないのであれば、そうじゃないやり方もあったりすると思って、どこまで広げるかっていう話は多分そのあたりがあるのかな。多分やる気を持つって思ってもなかなか難しいと思うので、そうすると、そのロジックモデルを作るプロセスの中で、今まで話ししたことのない人と話をすることで、その後仕事がしやすくなったりみたいなことが起きたりするっていう可能性も考えるファシリテーションが必要というか・・・極端な話、自分の仕事が増えましたみたいになる、要は何かやりたいこととかやった方がいいことっていう理解より、1タスクみたいな理解になってしまふと、難しい。だけど何かやることによって他の人と知り合うことができて、それで日常業務が円滑に進む。その環境を使ってロ

ジックモデルづくりに合うようなことが進んでいくかもしれないぞってちょっとと思える。そんな光がちょっと見えるんですが・・・。

E：（組織の）トップダウンで（参加・協働型評価を）制度化する。上を押さえるっていうのは結構重要だと思います。上の人には理解してもらうっていうのは大事だと思うんですけども仮にそこまでいかなくても・・・私の経験で思ったのは、5年ぐらい〇〇に直接関わってきましたけど、最初は本当に拒否ですよね。やっぱりね。また仕事が増えるから、でも変わった瞬間っていうのはね、確かにあって。この瞬間ってこれ、この「瞬間」って意味じゃないけど、この変わった時期っていうのがすごいわかるのは、ロジックモデルを使って自分たちの考え方を自分の政策を作るのに、自分たちが主体となって、関わる。関わることの面白さっていうのを感じた時。これ面白いって、正直みんな言ってました。それが自分たちの仕事の一部、私、仕事だから面白いんだと思うんですよね。そうなったときすごい変わりましたね。ただそういうのを支援するやっぱり伴走支援的な役割って結構必要かなと思います。

協働・参加型の評価における評価者の位置づけについては、評価実践や研究の発展の歴史では、第三者や外部評価者が関わることが多い。しかし、何を評価の対象とするのかということや、評価の範囲（バウンダリー）の設定、評価の目的にもよるが、内部者がファシリテーションを行う例はあり、内部者であるからファシリテーションができないというものでもない。ただ評価ファシリテーターに求められることは、その場にいる人が決められる、価値判断ができる事をいかにサポートするか、その場にいる人たちがコミットメントをもってそこに参加できるように働きかけるかであり、判断に参加者が迷うときに、評価的知識や思考でもってガイドする存在だと言える。このようなことを考慮すると、評価ファシリテーションガイドについては、ファシリテーションの実際に加えて、評価者の前提についての整理を明示することも必要であることが学びとして浮かび上がった。

6. ワークショップ後の発展と今後に向けて

2023年 のガイドの初版の実装後、フォーカスグループインタビューで得られた意見等を参考に、ガイドのバージョンアップを続けている。フォーカスグループインタビューで語られた参加者や企画者の経験からの語りは、ガイドのバージョンアップにもたらす知見が多かった。例えば、現在のバージョンでは、評価ファシリテーションワークショップの「場」に焦点化した内容となっており、その前段階における準備については、ファシリテーションをスムーズに進めて行くためのロジスティックな内容にとどまっていた。しかし、前段に示したように、「評価者」という前提の整理をしたうえで、立場性を考慮した具体的な内容を示す必要があるということを理解する助けにもなった。また、評価的な活動として、ロジックモデルが完成することも重要ではあるが、ロジックモデルを作成するというプロセスにおける対話が、ともにその場にいるメンバー同士の理解や関係性の深化に作用し、評価的な活動が結果として参加者の意識の変化につながるという効果についても言及があり、改めてファシリテーションの効用とそれをどのように評価のプロセスに入れ込んでいくのかを整理する点で示唆を得た。

これまで作成した私たちのガイドの内容を相対化し、より充実したものとするために、その後文献研究等を進める中で、いくつかの参考に値する資料評価ファシリテーションを進めて行くにあたっての貴重な研究と出会うことができた。例えば Patton (2018) が評価ファシリテーションの実践上の 5 つの原則 (Patton, 2018) を提言していた。その内容として①参加者個人の特徴を良くつかむ、②選択肢を提示する、③観察して、解釈して、適応する、④プロセスに評価的思考を埋め込む、⑤最先端の知見で活力を注入することが重要な原則として挙げられている。

① 「参加者個人の特徴を良くつかむ」ということは、ファシリテーションを行う対象者について理解し、相互に信頼し尊重し合う関係を築くことが重要で、評価は、参加者の知識、興味、動機、懸念などに合わせてカスタマイズすべきであるというものである。

② 「選択肢を提示する」は、ファシリテーションは、単に参加者が事前に決定された結果を形式的に承認するだけのプロセスでなく、主要な利害関係者を

対象とした評価ファシリテーションの目的は、可能性を模索し、選択肢を検討し、優先順位を確立し、意思決定を行うこととされる。評価ファシリテーターは、選択肢を特定したり、明確化したりすることを支援する。

③「観察して、解釈して、適応する」は、ファシリテーションはグループの徹底的な観察によって推進し、観察したすべてが「データ」であるというマインドを持ち、関わるというものである。そしてグループダイナミクスと優先目標への進捗状況をモニタリングして、ペースの調整、プロセスの変更、およびグループが望ましい結果に向かって前進し続けるための情報を提供する。

④「プロセスに評価的思考を埋め込む」は、評価的思考があることが、他のファシリテーションとの違いであり、すべての評価ファシリテーションがその思考や評価的な能力を高めていくまでのトレーニングという考え方である。

⑤「最先端の知見で活力を注入する」は、評価的思考や評価的能力は評価の基本をカバーするが、新しい考え方や創造的な機会等について、あらかじめ参加者に伝えるのも役割であることを指摘するものである。

このほかにも、Stevahn & King (2016) が利害関係者を巻き込みながら評価活動を行う際の 13 のステップをロジックモデルづくりに展開した例などがある。

前掲した通り、日本の評価実践領域においては多くが accidental な評価者であることを考えると、この原則のすべてを網羅しながら評価活動をファシリテートできる実践者を増やしていくことは簡単なことではない。しかし、すでに企画者や参加者の問題意識や実践知と重なる部分が認められ、すでに記載している内容は原則とも重なることが確認できた。今後、バージョンアップを進めて行く中で、これまで触れられていなかった内容等に関しては、ガイドの実装機会や練習機会の在り様とあいまって、引き続き共同研究者とともに検討ていきたい。

謝辞

本論文は 2023 年度しあわせ研究費（研究テーマ：評価ファシリテーションのモデル化と実装研究）の助成を受けたものです。また本稿執筆に係る研究において共同研究者であり、ガイドの共同執筆者である明治大学の源由理子先生、

日本社会事業大学の新藤健太先生、そして共同研究者でワークショップを共催頂いた川崎医療福祉大学の大山早紀子先生、立教大学の小野田由実子先生に大変お世話になりました。また評価ファシリテーションガイドの実装にあたり、研修に参加いただいた受講生の皆さんならびにフォーカスグループインタビューのご協力いただいた皆さんに感謝を申し上げます。

注釈

- 1 日本では特定非営利活動法人日本評価学会が「評価士養成講座」を開設し、講座修了と認定試験合格をもって「日本評価学会認定評価士」が付与される（日本評価学会, 2025）。
- 2 プログラム評価の「プログラム」とは、「何らかの課題解決」や「社会的価値の創出」に貢献するために設計された様々な取り組み（プロジェクト、事業、施策などと呼ばれる）=社会的介入を指す。源（2022）は、プログラム評価の主な要素として、①社会調査手法を活用し、データ収集・分析を行うこと、②プログラムの有用性を体系的に調査すること、③データとともに、何らかの基準と比較しながら価値判断が行われること、④プログラムを取り巻く組織的・政策的な文脈を考慮して行われること、⑤社会状況を改善するための活動に資する情報を提供することの5つを挙げている。

参考文献

- 池田葉月（2021）『自治体に評価における実用重視評価の可能性－評価結果の報告方法と評価への参加に着目して－』晃洋書房。
- 小田理一郎（2023）「訳者による序文」、アダム・カヘン著・小田理一郎訳『ファシリテーション－5つの在り方で場を見極め、10の行動で流れを促す』英治出版。
- 見田宗介（1966）『価値意識の理論－欲望と道徳の社会学』弘文堂。
- 田中博（2010）「フィリピンでのNGO教育事業参加型評価～ファシリテーターの役割とステークホルダーのエンパワメント」、『日本評価研究』、10（1）：75-93。

- 日本評価学会 (2025) 『評価士養成講座』 <https://evaluationjp.org/activities/course-certification/evaluator-training-course>
- 藤川吉美 (1994) 『規範科学の基礎－認識論的・価値論的・方法論的基礎－』 成文堂 .
- 源由理子編 (2016) 『参加型評価－改善と変革のための評価実践』、晃洋書房 .
- 源由理子 (2020) 「評価の 5 階層とプログラムセオリー」山谷清志監修 源由理子・大島巖編著 『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』 晃洋書房 . 31-49.
- American Evaluation Association. (2018). *The 2018 AEA Evaluator competencies*. <https://www.eval.org/Portals/0/Docs/AEA%20Evaluator%20Competencies.pdf>
- American Evaluation Association. (n.d.). *Values of the American Evaluation Association (AEA)*. <https://www.eval.org/About/About-AEA#:~:text=AEA%20has%20over%205%2C000%20members,more%20than%2080%20foreign%20countries>.
- Cousins, J. B., & Chouinard, J. A. (2012). *Participatory Evaluation Up Close: An Integration of Research-Based Knowledge*. Information Age Publishing.
- Cousins, J. B. ed. (2020). *Collaborative Approaches to Evaluation: Principles in Use*. Sage.
- Davidson, E. J. (2005). *Evaluation Methodology Basics: The Nuts and Bolts of Sound Evaluation*. Thousand Oaks, CA: Sage.
- Fierro, R. S., Schwartz, A., & Smart, D. H. (2016). *Editors' note*. In R. S. Fierro, A. Schwartz, & D. H. Smart. (Eds.), *Evaluation and Facilitation, New Directions for Evaluation*, 149, 7-17.
- Graeber, D. (2001). *Toward an anthropological theory of value: The false coin of our own dreams*, Palgrave Macmillan. (= 藤倉達郎訳、2022 『価値論－人類学からの総合的視座の構築』 以文社.)
- Liamputtong, P. (2020). Qualitative research methods, Fifth edition. (= 木原雅子・木原正博訳、2022 『質的研究法：その理論と方法－健康・社会科学分野における展開と展望－』 メディカル・サイエンス・インターナショナル.)

- Patton, M. Q. (2008). *Utilization-focused Evaluation* (4th ed.). Thousand Oaks, CA: Sage.
- Patton, M. Q. (2018). *Facilitating Evaluation*. SAGE.
- Podems, D. R. (2019). *Being an Evaluator: Your Practical Guide to Evaluation*. Guilford Press.
- Schwandt, T. A. (2015). *Evaluation Foundation Revisited: Cultivating a Life of the Mind of Practice*. Stanford, CA: Stanford University Press.
- Scriven, M. (1991) *Evaluation Thesaurus* (4th ed.). Newbury Park, CA: Sage.
- Shadish, W.R., Cook, T.D., & Leviton, L.C. (1991). *Foundations of Program Evaluation: Theories of Practice*. Newbury Park, CA: Sage Publications.
- Stevahn, L., & King, J. A. (2016). *Facilitating interactive evaluation practice: Engaging stakeholders constructively*. In R. S. Fierro, A. Schwartz, & D. H. Smart. (Eds.), *Evaluation and Facilitation, New Directions for Evaluation*, 149, 67-80.
- Weiss, C. H. (1998). *Evaluation: Methods for Studying Programs and Policies*. 2nd Edition, Upper Saddle River, NJ: Prentice Hall. (=佐々木亮監修 前川美湖・池田満監訳、2014『入門評価学—政策・プログラム研究の方法』日本評論社.)
- Wingate, L. A., Boyce, A. S., Becho, L. W., & Robertson, K. N. (2025) *Core Concepts in Evaluation: Classic Writings and Contemporary Commentary*. Sage.